Nonlinear Electrodynamics in Biological Systems

Nonlinear Electrodynamics in Biological Systems

Edited by

W. Ross Adey
Jerry L. Pettis Memorial Veterans Hospital
Loma Linda, California

and

Albert F. Lawrence
Hughes Aircraft Company
Carlsbad, California

PLENUM PRESS • NEW YORK AND LONDON

Library of Congress Cataloging in Publication Data

International Conference on Nonlinear Electrodynamics in Biological Systems (1983:
Loma Linda, Calif.)
Nonlinear electrodynamics in biological systems.

Proceedings of an International Conference on Nonlinear Electrodynamics in Biological Systems, held June 5-9, 1983, in Loma Linda, Calif.
Bibliography: p.
Includes index.
1. Electrophysiology—Congresses. 2. Electrodynamics—Congresses. 3. Nonlinear theories—Congresses. I. Adey, W. Ross. II. Lawrence, Albert F. III. Title.
QP341.I58 1983 574.8 84-11715

ISBN-13: 978-1-4612-9720-8 e-ISBN-13: 978-1-4613-2789-9
DOI: 10.1007/978-1-4613-2789-9

Proceedings of an International Conference on Nonlinear Electrodynamics in
Biological Systems, held June 5-9, 1983, in Loma Linda, California

©1984 Plenum Press, New York
Softcover reprint of the hardcover 1st edition 1984

A Division of Plenum Publishing Corporation
233 Spring Street, New York, N.Y. 10013

All rights reserved

No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the Publisher

PREFACE

The past half century has seen an extraordinary growth in the fields of cellular and molecular biology. From simple morphological concepts of cells as the essential units of living matter there has been an ever-sharper focus on functional organization of living systems, with emphasis on molecular dynamics. Thus, life forms have come to be defined increasingly in terms of metabolism, growth, reproduction and responses to environmental perturbations. Since these properties occur in varying degrees in systems below the level of cellular organization, there has been a blurring of older models that restricted the concepts of life to cellular systems. At the same time, a search has begun for elemental aspects of molecular and atomic behavior that might better define properties common to all life forms.

This search has led to an examination of nonlinear behavior in biological macromolecules, whether in response to electrical or chemical stimulation, for example, or as a means of signaling along a molecular chain, or as a means of energy transfer. Experimental knowledge in this area has grown rapidly in the past decade, and in some respects has outstripped theoretical models adequate to explain these new observations. Nevertheless, it can be claimed that there is now an impressive body of experiments implicating nonlinear, nonequilibrium processes as fundamental steps in sequential operations of biological systems. Many of these key findings have been independently replicated, but acceptance of these concepts in a biological milieu steeped in the traditions of chemical equilibrium phenomena has come slowly. Despite experimental rigor, independent replications and an increasing armamentarium of relevant physical models based on dissipative and dispersive processes, there are still those who describe these findings as "sensational," or "attributable to experimental error," or "still a long way from an adequate theoretical treatment."

We believe that the forward march of science demands a more positive approach. This Conference was convened to develop a forum for biologists, physical scientists and engineers. In an historical perspective, its timing occurred at a unique moment. For the first time, biological and physical scientists confront the same essential questions in the physics of matter. At issue for the physical scientist is the question of randomness in the physics of condensed matter; in the dynamics of molecules and in the interaction of molecules and electromagnetic fields. Theoretical constructs more or less adequate for the description of matter in the bulk have been shown to be inadequate to describe matter on the submicron level. A particular case in point is the ubiquity of solitary waves in both dynamical and static systems ranging from long chain molecules to chemically reacting solutions. The existence of such waves depends upon nonlinear energy exchanges between distinct modes, as for example the interaction between vibrational and exciton modes in long chain molecules.

To the biologist, now aware of highly nonlinear phenomena in cell membrane transduction of hormonal, neurohumoral and immunological stimuli, for example, these models from the frontiers of the physics of matter offer unparalleled opportunity to move beyond accepted equilibrium models of bioenergetics in cellular and subcellular systems. We may anticipate description in physical, rather than chemical terms of the essential steps in enzyme activation and inactivation, and in so-called "active" cellular transport mechanisms, for example.

For the engineer, there is in prospect the development of totally new computing devices and systems, operating at the molecular level. Logical operations and information storage would occur in biomolecules arranged in sheets or monomolecular films, with packing densities and array speeds orders of magnitude greater than is now possible. These developments will in large part, make use of the unique properties exhibited by matter as it is organized in living organisms.

What does the future portend? In broad perspective, recognizing the fallibility of speculation, we may look towards a physical, rather than a chemical basis in understanding biological processes. A physical understanding will not replace the elegant beauty of the chemical edifice. Rather, it will extend in exquisite detail the concepts of atomic and molecular behavior that have been the chemist's immense contributions to the biology of the twentieth century. In the biomedical sciences, this new knowledge may already have taken us to the threshold of essentially new concepts in cellular immunology, aging and cancer, as well as in our understanding of the organization of the brain, and of the ways in which brain cells "whisper together" in learning and recall.

We are deeply indebted to the Federal agencies and their administrative scientists who encouraged us to conduct this Conference; to the FDA Bureau of Radiological Health, the US Air Force Office of Scientific Research, the Office of Naval Research and the Veterans Administration. Their generosity permitted a unique international meeting that held the unflagging interests of a highly interdisciplinary audience. It is our hope that readers of this volume will also capture this atmosphere of excitement that pervaded the proceedings.

We gratefully acknowledge the support of the many staff and volunteers of the Jerry L. Pettis Memorial Veterans Hospital under the direction of Mr. R. Quinto and Mrs. B. Kruggel. The excellent conference arrangements bespeak their selfless efforts over many months. We are most grateful to Ms. J. Merrill for editorial assistance in preparation of this volume.

Finally, the second co-editor wishes to acknowledge the generous support provided by Office of Naval Research Contract Number N000 1483-C-0010 during the period of preparation for these proceedings.

H. Ross Adey
A.F. Lawrence

CONTENTS

NONLINEAR EFFECTS OF ELECTROMAGNETIC FIELDS ON WHOLE ORGANISMS, LIVING TISSUES AND TISSUE PREPARATIONS

Nonlinear, Nonequilibrium Aspects of Electromagnetic Field Interactions at Cell Membranes 3
W. Ross Adey

Use of Bone Cell Hormone Response Systems to Investigate Bioelectromagnetic Effects on Membranes *In Vitro* 23
Richard A. Luben and Christopher D. Cain

Strong Interactions of Radiofrequency Fields With Nucleic Acid 35
Mays L. Swicord, Glenn S. Edwards, and Christopher C. Davis

Nonthermal Resonant Action of Millimeter Microwaves on Yeast Growth 59
F. Keilmann and W. Grundler

Effects on Nervous Tissue of Exposure to Electromagnetic Fields 65
Hans-Arne Hansson

Neurotransmitter Release Stimulated in a Clonal Nerve Cell Line by Low Intensity Pulsed Magnetic Fields 79
G. Rein and R. Dixey

Natural AC Electric Fields In and About Cells 87
Herbert A. Pohl

NONLINEAR DYNAMICS OF ORGANIC MOLECULES, INCLUDING BIOMOLECULES

The Energetics and Molecular Dynamics of the Proton Pumping Photocycle in Bacteriorhodopsin 107
Robert R. Birge, Albert F. Lawrence, Thomas M. Cooper, Craig T. Martin, David F. Blair, and Sunney I. Chan

Presence of Nonlinear Excitations in DNA Structure and Their Relationship to DNA Premelting and to Drug Intercalation 121
Asok Banerjee and Henry M. Sobell

Solitons and Bioenergetics 133
Alwyn C. Scott

Nonlinear Dynamics of Globular Proteins 143
Peter S. Lomdahl

Solitons in Synthetic and Biological Polymers 155
A.R. Bishop

Picosecond Relaxations in Proteins and Biopolymers Observed by MM-Wave Spectroscopy 177
F. Kremer, A. Poglitsch, and L. Genzel

Nonlinear Mechanisms in Molecular Energy Transfer 187
T.W. Barrett

Nonlinear Excitations in Polyacetylene 201
W.P. Su

PROSPECTS FOR A BIOELECTRONIC TECHNOLOGY

Communication With Submicron Structures: Perspectives in the Application of Biomolecules to Computer Technology 207
Albert F. Lawrence and Robert R. Birge

Langmuir-Blodgett Films in Microelectronics 219
E.D. Thompson

Nonlinear Tunneling Barriers at High Frequencies and Their Possible Logic Processing Function in Biological Membrane 227
Gerald C. Huth, James D. Bond, and Per Arne Tove

Molecular Electronics: An Opportunity for a Biotechnical Synergism 243
Forrest L. Carter

NONLINEAR ELECTRODYNAMICS IN CELLULAR ELECTROCHEMISTRY

Radiofrequency Radiation and Chemical Reaction Dynamics . . . 277
Richard A. Albanese and Earl L. Bell

On the Growth of Cellular Tissues Under Constant and Environmental Conditions 287
Rene Lefever and Thomas Erneaux

Ionic Currents and Field Effects in Neural Extracellular Spaces 307
T. Triffet and H.S. Green

Frequency Selective Propagation of Extracellular Electrical Stimuli to Intracellular Compartments 327
H.P. Schwan

Electromagnetic Surface Wave Penetration of Biological Media 339
Leonard S. Taylor

MODELS OF NONLINEAR PROCESSES

Conceptual Models of Cooperative Processes in Nonlinear Electrodynamics 347
R.K. Bullough

Entrainment-Quasiperiodicity-Chaos-Collapse: Bifurcation Routes of Externally Driven Self-Sustained Oscillating Systems 393
Friedemann Kaiser

Field Induced Resonance and Chaos in Nonlinear Conservative Systems 413
L.E. Reichl

Solitary Waves As Fixed Points of Infinite Dimensional Maps 433
David W. McLaughlin, J.V. Moloney, and A.C. Newell

Nonlinear Phenomena in a Dispersive System With Application to Water Waves 443
Henry C. Yuen

Boson Condensation in Biological Systems 469
G. Vitiello and E. Del Giudice

Order and Structures in Living Systems 477
E. Del Giudice, S. Doglia, and M. Milani

APPLICATIONS OF NONLINEAR PHYSICS IN BIOPHYSICS, BIOCHEMISTRY AND CELL BIOLOGY

General Theory of Coherent Excitations in Biological Systems 491
H. Frohlich

Millimeter-Wave and Far-Infrared Spectroscopic Properties of Cellular Systems 497
K.H. Illinger

Spectroscopic Studies of Nerves 519
B. Simic-Glavaski

The Modification of Davydov Solitons by the Extrinsic H-N-C=O Group 531
Scott P. Layne

Nonlinear Phenomena in Bioenergetics and Oncology as Seen in 25 Years of Research With Millimeter Microwaves and Raman Spectroscopy 549
Sydney J. Webb

Nonlinear Electrodynamics in Cytoskeletal Protein Lattices 567
Stuart R. Hameroff, Steven A. Smith, and Richard C. Watt

Bridging the Gap 585
Samuel Koslov

Index 599

NONLINEAR EFFECTS OF ELECTROMAGNETIC FIELDS ON WHOLE ORGANISMS, LIVING TISSUES AND TISSUE PREPARATIONS

NONLINEAR, NONEQUILIBRIUM ASPECTS OF ELECTROMAGNETIC FIELD INTERACTIONS AT CELL MEMBRANES

W.R. Adey

Research Service, Pettis Memorial Veterans Hospital and
Departments of Surgery and Physiology Loma Linda
University, School of Medicine Loma Linda, California
92357 USA

1. Introduction

Early concepts of the cell as a unitary biological entity postulated a role for the cell membrane simply as a physical boundary enclosing the intracellular contents. At the turn of the century, awareness of differential concentrations of potassium and sodium ions inside and outside this boundary focused attention on its capacity to sustain these differentials by expenditure of metabolic energy, in accordance with equilibrium considerations. Action potentials in excitable cells meet all criteria for transmembrane currents based on transient collapse and reestablishment of these ionic equilibria. In this model, the lipid bilayer or plasma membrane functions as the essential structural element separating intracellular contents from surrounding fluid.

It is now clear that cell membranes function in a far more complex fashion in detection of a wide range of external chemical stimuli, including neurotransmitters, antibodies and most hormones. Recent research has focused on the arrangement of proteinaceous material on and within the lipid bilayer, as described in the fluid-mosaic model of Singer and Nicolson (1). Glycoprotein (glycosaminoglycan) strands protrude outward from intramembranous particles (IMPs) to form a polyanionic surface layer composed of sialic acid terminals. Since this surface sheet contains receptor sites for hormones, antibodies and neurotransmitter molecules (2), it is the site of initial binding events for humoral stimuli bathing the cell. These first events at surface receptors are signaled to the cell interior, where intracellular responses are triggered at energy levels far higher than in the initial interactions at the cell surface, or in the energy of the electromagnetic (EM) fields now known to modify these intracellular responses to molecular stimuli at the membrane surface.

Although these basic steps in transductive coupling at cell membranes are now generally accepted, this quite simple concept is based on a series of observations requiring closer scrutiny if the essential elements in the structural and energetic sequence of transductive coupling are to be understood. Our studies have shown that cell responses to many natural humoral stimuli in the surrounding fluid can be manipulated either by certain EM fields at low frequencies, or by radio frequency and microwave fields amplitude modulated at low frequencies. These effects occur at field levels which are "athermal" in terms of tissue heating; and the field amplitude and frequency characteristics offer clues to the essential nature of membrane transductive coupling.

This model of transductive coupling at low imposed field levels has defined requirements for "amplifying" mechanisms between initial cell surface events and responses in intracellular enzymatic and metabolic systems. Two major energy barriers intervene between initial weak events at the membrane surface and ensuing major intracellular events. One is the barrier of the membrane potential that exists across the 40 A of the plasma membrane, with a typical transmembrane gradient of 10^5 V/cm; the second arises in the relative energies of molecular binding at membrane receptor sites and in intracellular events initiated by these binding events. As discussed below, an understanding of the striking cellular effects of oscillating EM fields in pericellular fluid many orders of magnitude weaker than the transmembrane gradient requires a membrane model based on "amplification" of these weak cell surface events at receptor sites. In the energetic barrier, intracellular enzyme systems, including adenylate cyclase and protein kinases, respond to imposed weak EM fields by alterations in energy release many orders of magnitude greater than in either initial binding events at cell surface receptor sites, or in EM fields that modify these intracellular responses to molecular stimuli at the membrane surface.

Of the neurohumoral, endocrine and immunological reactions at cell membranes so far examined, many display "windows" in their sensitivity to frequency and amplitude of the imposed fields. New evidence indicates that temporal windows also delimit field-sensitive epochs in modulation of intracellular enzyme systems by signals arising in cell membranes. Morphological data also support a sequential model of transductive coupling of EM field stimuli from cell surface receptor sites to the cytoskeleton via transmembrane macromolecules. These macromolecules also act as coupling proteins between surface receptor sites and enzymes at the internal membrane surface.

A clear emergent conclusion from these studies concerns the highly nonlinear character of transmembrane coupling of many cell surface stimuli, and the importance of associated nonequilibrium events. From many experimental and theoretical findings cited below, a substantially new view is rapidly evolving about mechanisms for signaling from the cell surface to the interior, and their possible role in an intimate communication system between cells in tissue, by which they may "whisper together" (Young, 1951).

In summary, there is good evidence from studies described below that coupling of weak, oscillating electrochemical fields occurs at membrane surface polyanions (formed by the protruding glycoprotein terminal strands on intramembranous macromolecules) and that these processes are Ca-dependent (3-8). The transductive coupling of many humoral stimuli at cell surface sites implicates a minimal sequence of three steps (4, 5, 9). Initial events appear to relate to the numerous negative charges at terminals on the surface glycoprotein sheet, a step presumed to occur in length and area of the membrane; a second step involving transmembrane signaling, mediated by lipoproteins that span the transverse dimension of the plasma membrane; and a third step coupling the transmembrane signal to intracellular organelles via cytoskeletal elements. We have shown that imposed EM fields interact with at least the first and second steps of this sequence.

Is there evidence that imposed EM fields at levels no greater than those seen in intrinsic electrochemical oscillations in tissue can perturb ionic mechanisms controlling cellular excitability? Our first studies examined EM field effects on Ca efflux from brain tissue (10-15), resulting in clear evidence of highly cooperative interactions. Is there evidence that Ca-dependent membrane processes, including release of neurotransmitters, responses to hormones and cell-mediated immune mechanisms, are modulated by the same types of EM fields? As reviewed below, there is impressive and growing evidence consistent with nonlinear, nonequilibrium interactions in all three types of cell mechanisms, with responses in widely diverse tissues that include brain, bone, pancreas and leucocytes.

These data provide a reasonable basis for modeling many of the major steps in the sequence of cell membrane transductive coupling; beginning with events at binding sites that involve modulation of Ca binding, followed by modulation of membrane-bound enzyme activation, as in the case of adenylate cyclase (16). As a consequence of these and other membrane interactions, there is modified activity in the great class of intracellular messengers, the protein kinases.

2. Actions of EM fields on Ca binding and transmitter release in brain and other excitable tissues

a. Ca efflux from cat cerebral cortex in response to weak electrical stimuli

Studies in the awake cat (10) showed binding and release of Ca in cerebral cortex to be sensitive to electrical stimuli of the same amplitude as the gradient of the EEG at cellular dimensions. Pulsed stimuli (200/sec, 1.0 msec duration) under conditions of relatively uniform fields of 50 to 100 mV/cm increased $^{45}Ca^{2+}$ efflux by almost 20 per cent. A concomitant rise in efflux of the amino acid transmitter GABA is described below (Fig. 1).

b. Ca efflux from cat cerebral cortex exposed to modulated microwave fields

We have studied $^{45}Ca^{2+}$ efflux from awake cat cerebral cortex during and following exposure for 60 min to a 3.0 mW/cm^2 450 MHz field sinusoidally amplitude modulated at 16 Hz (14). We used the same superfusion techniques as in the electrical stimulating experiments, employing a plastic cylinder in gentle contact with the pial surface. Tissue dosimetry showed a field of 33 V/m in the interhemispheric fissure (rate of energy deposition 0.29 W/kg). Radioactivity measurements in cortical samples after superfusion showed $^{45}Ca^{2+}$ penetration at about 1.7 mm/hr, consistent with diffusion of the ion in free solution. By comparison with controls, efflux curves from field-exposed brains were disrupted by waves of increased $^{45}Ca^{2+}$ efflux. These waves were irregular in amplitude and duration, but many exhibited periodicities of 20-30 min. They continued into the postexposure period (Fig. 2). Binomial probability analysis indicated that these field-exposed efflux curves constitute a different population from controls at a confidence level of 0.96.

In about 70 percent of cases, initiation of field exposure was followed by increased end-tidal CO_2 excretion for about 5 min. However, hypercapnea induced by hypoventilation did not elicit increased $^{45}Ca^{2+}$ efflux. Thus, this increased efflux with exposure did not appear to arise as a secondary effect of raised cerebral CO_2 levels. However, efflux rates may depend on H^+ ion levels, since removing HCO_3^- ions from the medium bathing isolated chick cerebral cortex reduced $^{45}Ca^{2+}$ efflux in both control and exposed preparations and raising H^+ ion levels in the presence of HCO_3^- ions increased efflux sensitivity to field exposure (13).

These interactions between H^+ and Ca^{2+} ions suggest a competition for binding sites at cell membrane polyanions. The Ca efflux experiments also indicate a role for these ions in excitatory processes, as proposed by Adey et al. (14).

c. Ca efflux rates from cerebral synaptosomes exposed to modulated microwave fields

So far, we have considered evidence indicating sharply altered Ca binding in cerebral tissue in response to imposed electric fields at levels mimicking natural extracellular electrochemical oscillations. Further studies have defined similar sensitivities in preparations of ultrastructural cerebral elements. Cerebral synaptosomes, typically 0.7 m in diameter, retain the synaptic junction and adjoining postsynaptic membrane and have characteristics typical of brain chemical synapses. Lin-Liu and Adey (15) studied Ca efflux from pellets of synaptosomes preloaded with $^{45}Ca^{2+}$, employing a continuous perfusion technique in a Ca-free physiological solution. This minimized Ca-Ca exchange between intra- and extracellular compartments (15). A 450 MHz field at a power density of 0.5 mW/cm^2 (measured gradient in air, 43 V/m) and sinusoidally modulated at 16 Hz increased the rate constant of the efflux by 38 percent. Unmodulated and 60 Hz modulated fields were without effect (Fig. 3). This study also supports a model of EM field interaction with Ca bound at cell membrane surface sites. This field-induced change in efflux was distinguishable from $CaCl_2$-stimulated efflux. The latter is most probably derived from intracellular sites.

d. Ca efflux studies in isolated chick cerebral hemisphere

The strongest single line of evidence about the essential nonlinearity of these bioeffects has come from manipulation of field frequencies and intensities. Much further study is needed to unravel mechanisms of "windowed" responses that occur, for example, only for narrow bands of frequencies and intensities in the spectrum below 100 Hz, or in multiple intensity windows (19). These windowed responses were initially reported as modifications of cerebral Ca efflux, but also occur in Ca efflux from cardiac muscle (20), in cell mediated cytolysis (21), in steroid hormone secretion from the isolated adrenal gland (22), and in the activation of protein kinases in human T-lymphocytes (17).

Ca efflux from cerebral cortex is sensitive to radiofrequency (147 MHz) and microwave (450 MHz) fields sinusoidally amplitude-modulated at frequencies below 35 Hz, at tissue intensities similar to the EEG measured across the dimensions of a single neuron (20-100 mV/cm). $^{45}Ca^{2+}$ efflux from isolated chick cerebral hemispheres increased more than 15 percent over control levels at a modulation frequency of 16 Hz, producing a striking "tuning curve", with progressively smaller responses at frequencies above and below 16 Hz (23, 24). This increased $^{45}Ca^{2+}$ efflux persisted in the presence of 10^{-4} M NaCN, precluding metablic activity and active Ca transport (Fig. 4).

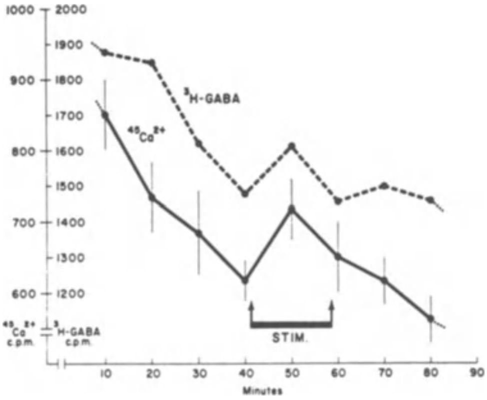


Fig. 1 Efflux ^{45}Ca^{2+} and ^{3}H-GABA from superfused cat cortex, with effect of weak electrical stimulation (200/sec, 1 msec pulses). Electrical gradients were of the order of 50 mV/cm (From Kaczmarek and Adey, 1974).

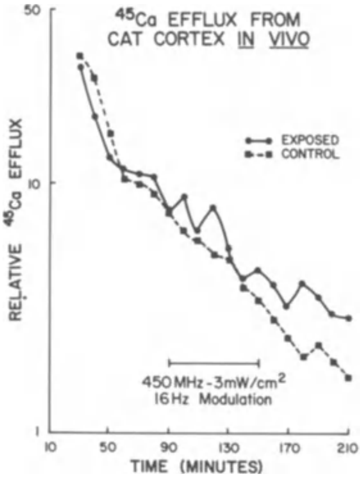


Fig. 2 Increased ^{45}CA^{2+} efflux from awake cat cerebral cortex during and following exposure for 1 h to a 3 mW/cm^{45}Ca^{2+} 450 MHz field, sinusoidally modulated at 16 Hz (From Adey, Bawin and Lawrence, 1982).

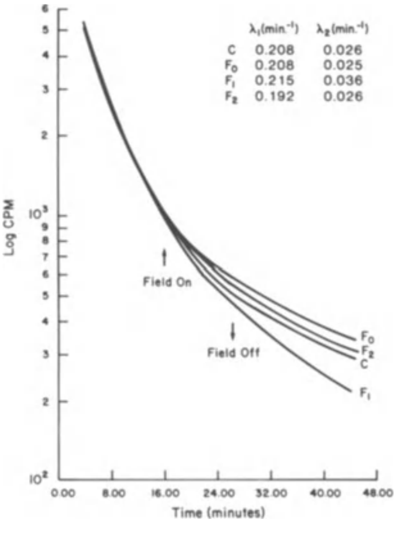


Fig. 3 Computer fitting of ^{45}Ca^{2+} efflux data composites from rat cerebral synaptosomes exposed to 450 MHz fields, 0.5 mW/cm^2, sinusoidally modulated at 16 Hz. C, control; F_0, unmodulated 450 MHz fields; F_1, same field with sinusoidaly 16 Hz modulation; F_2, same field with 60 Hz modulation (From Lin-Liu and Adey, 1982).

ELF electric fields in the range 1-100 Hz inducing far weaker tissue components (calculated at 10^{-7} V/cm) modified Ca efflux as a virtual mirror image of effects of modulated RF fields (Fig. 4). In the same range of low frequencies as those modulating the RF fields, $^{45}Ca^{2+}$ efflux from isolated chick cerebral hemisphere was reduced by about the same amount as it was increased in the RF fields, and maximal interaction again occurred around 16 Hz (11). By contrast, Ca efflux was increased when these weak electric fields were combined with an H field induced by current flowing synchronously with the electric field (25). Interpretation of Ca efflux experiments is difficult, since sharp differences in binding energy may exist within and between stranded macromolecules on the membrane surface (26).

Intensity windows in release of Ca from cerebral cortex have also been reported for both ELF and modulated RF/microwave fields. Blackman et al. (24) noted that 147 MHz fields sinusoidally modulated at 16 Hz were only effective in increasing Ca efflux from chick cerebral cortex at incident energies around 1.0 mW/cm^2. This was subsequently confirmed with 450 MHz fields modulated at 16 Hz, with an amplitude window for intensities between 0.1 and 1.0 mW/cm^2 (12). For ELF fields, Blackman et al. (27) observed intensity-dependent Ca efflux changes over the range of field intensities in air of 1-70 V/m for 16 Hz ELF fields.

For brain tissue, pulsed electric fields that impose pericellular gradients of the same amplitude as the EEG recorded across the dimensions of a single cell increase the efflux of the inhibitory transmitter amino acid GABA by about 20 percent (10). Other studies in cultured nerve cells (rat pheochromocytoma) exposed to a 500 Hz 8.5 G pulsed magnetic field (induced electric field around 10^{-4} V/cm) also showed a sharply increased release of the neurotransmitter nonadrenaline (8). Increasing the extracellular Mg concentration about 7 times did not alter the release rate in control studies but the stimulated release was abolished. These authors concluded that Mg^{2+} appeared to act as a specific Ca^{2+} antagonist; and that cation binding sites on the outer surface of the plasma membrane may be highly sensitive to weak EM stimuli. In their model ensuing cooperative changes in Ca^{2+} binding would modify membrane stability to allow Ca^{2+} entry through the membrane and release of transmitter vesicles.

These data offer strong evidence that brain tissue responds to intrinsic and imposed oscillating electric fields far weaker than the membrane potential by altered Ca binding at cell membrane surfaces. These responses are highly nonlinear with respect to field frequency and intensity. What, then, are the physiological concomitants of this altered Ca binding? Is this sensitivity to weak pericellular electric oscillations confined to brain tissue, or is it a widespread property of cells in other tissues?

We shall examine evidence from manipulation of Ca-dependent membrane activities in nonneural tissues. They also share this sensitivity to weak imposed EM fields.

3. EM field sensitivities in bone, pancreas and lymphocytes

Major steps in the sequence of transductive coupling at cell membranes have been revealed by their manipulation with EM fields. This picture has emerged in our collaborative studies from examination of widely divergent tissues and cell lines, including bone and bone cell cultures (16), cultured pancreatic islets (28), and in human and murine lymphocytes (17, 21). We have noted field modulation of activation of a membrane-bound enzyme (adenylate cyclase) by hormone molecules that attach to membrane surface receptor sites, but no field modulation of the same enzyme's response to humoral substances that act at the cell nucleus, rather than at the membrane. There is altered activity in intracellular messenger enzymes (protein kinases) as they respond to field-induced signals that appear to come directly from the cell membrane but not via membrane-bound enzymes. Field interactions also modify secretory capacities in these cells, including hormone release and hormonal control of collagen fiber deposition in the intercellular matrix. Allogeneic cytolysis of target cells is modulated by EM fields in a clear example of a membrane-membrane interaction through antigen-antibody reactions. Many of these effects exhibit "windowed" characteristics to EM field frequency and amplitude, and some have shown windowing in the timing of their sensitivity to field exposure; thus emphasizing their nonlinear, nonequilibrium character.

a. Modulation by low-frequency EM fields of responses of bone and bone cells to parathyroid hormone (PTH) and 1,25 dihydroxy-vitamin D_3

In pulsed EM field studies, we have focused on the membrane-bound enzyme adenylate cyclase and its sensitivities to two agents. One acts at the cell membrane and the other at intracellular sites, probably at the nucleus. Manipulation of these actions by concurrent EM field exposure has confirmed the cell membrane as a prime site of EM field transduction, and at the same time appears to preclude a direct field action on intracellular organelles, such as the nucleus, as direct participants in these particular responses. Moreover, each step in both these manipulations is Ca-dependent.

We have examined effects of two pulsed EM fields, identical with those used in clinical management of ununited fractures, on neonatal mouse cranial bones and on cultures of a clonal line of osteoblast-like mouse bone cells (16). One field generated continuous pulse trains at 72 Hz and the other generated recurrent bursts (5 msec duration) of pulses at a 4kHz rate (burst repetition rate 15/sec). No differences were noted in responses to the two fields. Induced extracellular fields were 1 mV/cm or less, and the expected peak extracellular current density in a homogeneous conducting electrolyte was of the order of 1.0 $\mu A/cm^2$. Transmembrane components of extracellular field gradients were as much as six orders of magnitude less than typical excitatory threshold currents of 1 mA/cm^2 observed for axonal depolarization.

Following field exposure for 48 to 90 hours, tissue culture medium was replaced with fresh medium containing 0.01 mM theophylline. After 30 mim, hormones (either parathyroid hormone (PTH) or 1,25 dihydroxy-vitamin D_3) were added and assays performed for adenylate cyclase activity through cAMP accumulation and for collagen synthesis. Both fields significantly reduced cellular production of cAMP, even though neither basal nor fluoride-activated levels of adenylate cyclase were altered in membranes from cells cultured in the fields. The fields also blocked the inhibitory effects of PTH on collagen synthesis by these cells. However, there were no field effects on inhibition of collagen synthesis by vitamin D_3, which is believed to occur primarily by a nuclear, rather than by a membrane-dependent mechanism; thus supporting the hypothesis that field actions are primarily at cell membrane sites (Fig. 5).

b. Decreased protein kinase activity in human lymphocytes exposed to low-level, amplitude-modulated microwave fields

The foregoing study provided evidence that activation of membrane-bound enzymes in response to hormones bound at membrane surface receptor sites is extremely sensitive to weak, low frequency electrochemical oscillations in the pericellular environment. In this case, activation of adenylate cyclase converts adenosine triphosphate (ATP) to 3'5' cyclic adenosine monophosphate (cAMP). In turn, cAMP is responsible for activation of a large portion of a group of messenger enzymes, the protein kinases. Other protein kinases are activated by pathways from the cell membrane not involving cAMP. A possible mechanism by which hormone-receptor interactions at the cell surface lead to alterations in intracellular events is by the reversible phosphorylation of cellular proteins by protein kinases. For this reason, we have investigated the ability of brief exposure to weak microwave fields sinusoidally modulated at low frequencies to alter protein kinase activity of human tonsil lymphocytes in culture (17). We have examined levels of cAMP-dependent and cAMP-independent protein kinases, thus distinguishing those kinases activated through the cAMP production path outlined above from those activated through signals via other paths from the cell membrane. The activation of protein kinases is generally Ca-dependent.

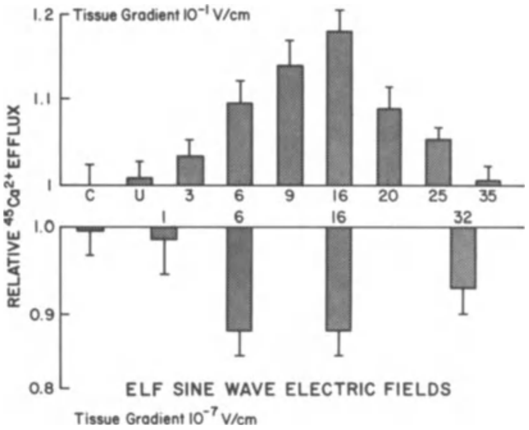


Fig. 4 A: ^{45}Ca^{2+} efflux from freshly isolated chick cerebral hemispheres exposed to a weak RF field (147 MHz, 0.8 mW/cm^{45}Ca^{2+}, amplitude modulated at low frequencies (abscissa). B: Effects of far weaker electric fields (56 V/m) in the same frequency spectrum from 1 to 32 Hz. Field gradients in tissue differ by about 10^6 between A and B (From Adey, 1981 b).

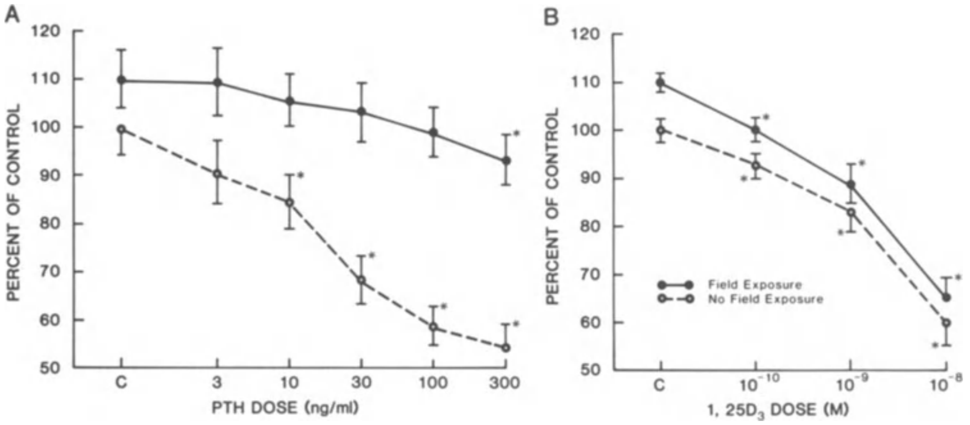


Fig. 5 Effects of 72 Hz pulsed magnetic field on synthesis of collagen by bone cells, grown in presence (solid line) or absence (dashed line) of field for 12 h. Field blocked inhibitory action of PTH (A) but exerted no effect on inhibition of collagen synthesis by vitamin D3 (From Luben et al., 1982).

Cultures were exposed for periods up to 60 min at 35C to a 450 MHz field (peak envelope power 1.5 mW/cm^2), amplitude modulated at frequencies from 3 to 100 Hz (modulation depth 80 percent). Protein kinase activity was estimated with calf-thymus histone as the substrate. Approximately 20 percent of the total kinase activity was due to specific activity of cAMP-dependent protein kinase, which was not altered by exposure to fields modulated at 16 or 60 Hz for 15, 30 or 60 min.

On the other hand, the total cAMP-independent protein kinase activity was sharply reduced by more than 50 percent 15 and 30 min after onset of exposure to a 16 Hz modulated field, but returned to control levels after 60 min, despite continued presence of the field (Fig. 6). The transient nature of this response was also seen with 60 Hz modulation, but was smaller than with 16 Hz modulation (20 percent at 15 minutes and 25 percent at 30 minutes). We have confirmed the existence of a "window" with respect to modulation frequency, with maximum effects at 16 Hz. Effects decrease progressively with 6 Hz & 3 Hz modulation and with an unmodulated field; and at higher modulation frequencies (40, 60, 80 and 100 Hz).

Thus, this study reveals not only highly nonlinear and nonequilibrium sensitivities as indicated by the modulation frequency window; it also indicates a brief window in time during which the lymphocyte is susceptible to this electrochemical oscillation. Manipulation of the steps in transductive coupling with imposed EM fields has confirmed the essential sequence. It would be expected that the phenomenon of temporal windowing may be of biomedical significance in definition of hazards attributed to human environmental exposures to EM fields and also in delimiting their therapeutic application.

c. Reduced cytotoxicity of allogeneic T-lymphocytes in modulated microwave fields

Cell mediated immune responses of allogeneic T-lymphocytes, involving cytolysis of target cells, are mediated at cell membranes and require the presence of Ca ions. We have tested this response as a function of modulation frequency of a weak 450 MHz field (incident energy 1.5 mW/cm^2 peak envelope power) (21) in a mouse line of cytotoxic T-lymphocytes (CTLL-1) targeted against human lymphoma cells (H_2^dB MPC-11).

Cytotoxicity was reduced about 20 percent when a 4 h assay was performed with 60 Hz sinusoidal modulation of the field. This sensitivity followed a "tuning curve" against modulation frequency, with peak suppression at 60 Hz and smaller effects at all other frequencies from 3 to 100 Hz (Fig. 7). A similar inhibition occurred when T-lymphocytes were exposed prior to mixing

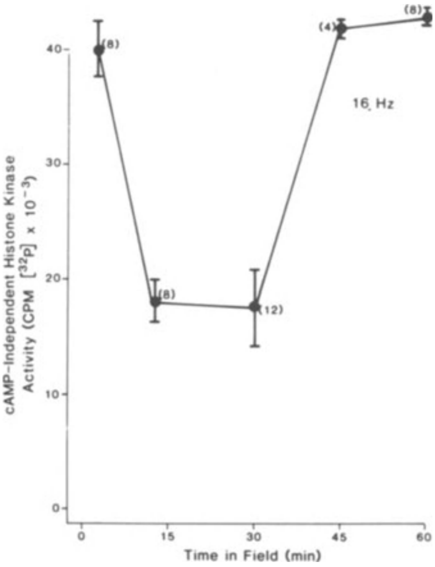


Fig. 6 Transient decrease in protein kinase activity at 15 and 30 min in human lymphocytes exposed to 450 MHz field (1.7 mW/cm^2 peak envelope power) sinusoidally modulated at 16 Hz (From Byus et al., 1983).

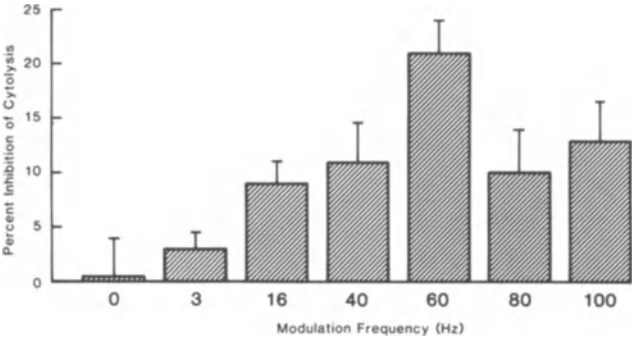


Fig. 7 Inhibition of cytotoxicity of allogeneic T-lymphocytes exposed to a 450 MHz field (1.7 mW/cm^2 peak envelope power) as sinusoidal amplitude modulation was varied between 0 and 100 Hz.

with the target cells, indicating their direct interaction with the field. Exposed T-lymphocytes recovered their full cytolytic capacity in 12.5 hours. Unmodulated fields were without effect.

d. Altered Ca binding and reduced insulin secretion in pancreatic islets exposed to low frequency pulsed magnetic fields

The foregoing studies have reviewed the sequence of steps in coupling various physiological stimuli from the cell surface to the interior and the electrochemical energy levels that appear to be involved at each major step. There is also evidence that the same EM fields influence secretory functions in response to humoral stimuli. Effects on collagen formation by bone cells were cited above (16).

Using a "pulse burst" type of EM field similar to that used in the bone cell study, Jolley et al. found that rabbit pancreatic islets showed a 26 percent reduction in Ca content, a 25 percent reduction in Ca efflux, and a 35 percent decrease in insulin release in response to glucose stimulation (28). A 3-fold increase in the steroidogenic response of rat adrenal tissue *in vitro* to 10 mU ACTH has been reported following exposure to a 60 Hz electric field at 10 kV/m. The study suggested a field intensity "window", with smaller reponses at 500 and 1000 kV/m. For an external field of 1000 kV/m, the field intensity in fluid surrounding the tissue was only 16.8 mV/m (22). The steroid response was measured as the corticosterone level in the superfusate 5.5 hours after field activation, and 1.5 hours after ACTH stimulation.

4. Nonlinear electrodynamics in cell membranes at the molecular level

We have reviewed the evidence that points strongly to the nonlinear, nonequilibrium behavior of transductive systems that couple a wide range of physiological stimuli from the cell surface to intracellular mechanisms. This new vista on cell membrane mechanisms results from the use of weak oscillating EM fields as concomitant manipulanda. It is clear that many of these observations have been duplicated, or complement each other in consistent ways. These imposed EM fields are orders of magnitude weaker than would have been predicted for threshold effects from classical models based on equilibrium considerations, including the membrane potential. There are at least three major considerations to be addressed in future research.

A basis must be found for the biological transduction of these electrochemical oscillations in pericellular fluid at frequencies below a few hundred Hz when they are so much weaker than the membrane potential. No existing models based on equilibrium considerations, such as the Hodgkin-Huxley model of excitation in nerve fibers, address this problem appropriately.

It is also necessary to elucidate the "windowed" character of many of these responses with respect to these low frequencies, and to take account of similar windowing in relation to intensity. We conclude that these interactions are nonequilibrium in character and also highly nonlinear (3-5).

There is the further problem of intercellular communication which may be mediated by weak electrochemical oscillations generated in one cell and transmitted to adjacent cells. So far, we have addressed membrane transduction from the exterior to the interior of the cell, since this is the direction of one major signal flux. On the other hand, low frequency oscillations in membrane potentials of cerebral neurons produce currents in pericellular fluid that influence adjacent cells (29, 30). Intercellular communication based on low frequency oscillations is a clear possibility. The extracellular medium is an efficient conductor at low frequencies and may behave in a frequency dependent manner in the counterion layer adjoining the membrane surface. Leakage of low frequency membrane potential oscillations into the extracellular space has been extensively described, and plays a prime role in brain tissue in the genesis of the EEG. (For review, see [4]).

At far higher frequencies, as for example, in the far infrared region, such a model presents considerable difficulties. Generators of coherent infrared radiation are not known to exist on cell surfaces nor in the intercellular space. On the other hand, these high frequency models, as proposed by Fröhlich (31), may be of great importance in the regulation of intracellular events. We may speculate that release of metabolic energy within the cell, as in the hydrolysis of ATP, may be associated with release of coherent infrared radiation. If this should prove to be the case, intracellular signaling and energy transfer might well be so mediated over short distances. How much of this coherent radiation might pass through the cell membrane to modulate activity of adjacent cells? An answer must await future research, but absorption and scattering within the lipid bilayer would appear to make this unlikely.

We have hypothesized a two stage process in transductive coupling from membrane surface to cell interior in an attempt to unify known elements in this sequence (9). We propose an initial interaction over whole areas of the membrane surface. This is dissipative in nature and highly cooperative in its modulation of Ca binding at sites on terminals of stranded glycoproteins. The ensuing step would be dispersive, with solitary waves similar to those proposed by Davydov moving in sequence down the length of glycoprotein and lipoprotein molecules. Since many of these molecules span the plasma membrane from its external surface to the cell interior, they form a path across cell membranes for signaling or for transport of energy, after initial transductive events first modify surface electrochemical states by cooperative interactions in the length and area of the membrane surface (3, 4, 9, 32, 33). Birge's model of bacteriorhodopsin may serve as a paradigm. Light energy is focused at a specific site, an electrostatic attraction is broken, and an energy or charge product passes to the cell interior by molecular conduction (34).

The extracellular space forms a preferred pathway for pericellular fields, directing induced currents along the cell surface. Initial transduction of weak EM fields would therefore occur in the plane of the membrane surface. We have proposed that this dissipative interaction rests upon formation of "patches" on the membrane surface. Within each patch, fixed charge sites on protruding glycoprotein strands achieve coherent energy levels substantially above ground states, and with identical energy levels at neighboring sites. Oscillating EM fields or proton tunneling at the boundaries of these coherent domains may act as weak triggers, initiating a domino effect across the entire domain and returning it to ground state with release of much more energy than in the initial triggering event. Altered Ca binding at surface sites induced by field exposure supports this model (4). This highly nonlinear modification of Ca binding may offer a key to the "amplifying" mechanism inherent in the observed EM field sensitivities of neurohumoral, immune and endocrine events initiated at cell surface sites. This amplification in the first stage of transductive coupling may provide energy for the second or transmembrane step in stimulus coupling.

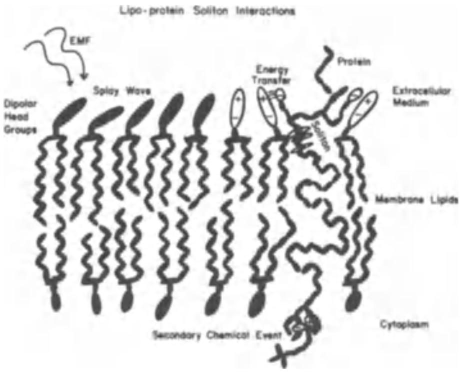


Fig. 8 Hypothetical interaction between an EM field-induced wave in the membrane and a membrane protein. The energy of a splay wave traveling along the membrane is transferred to a protein via coulombic coupling between dipole-type head groups and polar sites along the protein chain (From Lawrence and Adey, 1982).

In this model, linear macromolecules spanning the membrane play a key role in signal coupling to the cell interior. This coupling would be a direct consequence of initial longitudinal events on the membrane surface and may be mediated by soliton processes (9). Recent research has suggested that nonlinear interatomic forces (specifically in the hydrogen bond) can lead to robust solitary waves with greatly increased radiative life times (35, 36) (Fig. 8). Davydov (35) concludes that this would correspondingly increase the tendency for molecular vibrations to be the vehicle for energy transfer over long molecular chains, specifically over the "spines" in alpha-helix proteins with the bond sequence ---HNC=O---HNC=O-, etc. Davydov's nonlinear analysis shows that propagation of amide-I vibrations is coupled to longitudinal soundwaves in the alpha-helix, and the coupled excitation propagates as a localized and dynamically stable wave. The amide-I vibrations are the source of the longitudinal sound waves which stretch hydrogen bonds along the α-helix. Soliton formation exhibits a sharp threshold which is a function of energy coupling between the hydrogen bond stretching and the amide-I excitation (37). With a further increase in coupling energy, there is a second threshold above which solitons will not form. It may be asked whether the general properties of these nonlinear mechanisms can account for peristent molecular states in the needed time frame.

Much biochemistry has rested hitherto on such descriptive terms as "activation" of enzymes, or "active transport". This research may open the door to an understanding of these processes in physical terms. These concepts are as new to the physical sciences as in their applications to biology and medicine, and suggest a fundamentally new aspect to the organization of matter at the molecular atomic level, where there is a property of "graininess". Manifestations of these characteristics in biological molecules may ultimately offer an understanding of essential properties of living matter.

Acknowledgments

Studies from this laboratory were generously supported by the Veterans Administration, the U.S. Department of Energy, the Bureau of Radiological Health (FDA), and the Southern California Edison Company.

References

1. S.J. Singer and G.L. Nicolson, The fluid mosaic model of the structure of the cell membrane, Science 175:720 (1972).

2. R.K. Gershon, "Clonal selection and after," and after, New England J. Med. 300:1105 (1979).

3. W.R. Adey, Ionic nonequilibrium phenomena in tissue interactions with electromagnetic fields, *in*: "Biological Effects of Nonionizing Radiation," K.H. Illinger, ed. American Chemical Soc, Symp. Ser. 157. American Chem. Soc., Washington, D.C. (1981).

4. W.R. Adey, Tissue interactions with nonionizing electromagnetic fields, *Physiol. Rev.* *61*:435 (1981).

5. W.R. Adey, Biological effects of low energy electromagnetic fields on the central nervous system, *in*: "Biological Effects and Dosimetry of Nonionizing Radiation," M. Grandolfo, S.M. Michaelson and A. Rindi, eds. Plenum, NY (1983).

6. W.R. Adey, Some fundamental aspects of biological effects of extremely low frequency (ELF) radiation, *in*: "Biological Effects and Dosimetry of Nonionizing Radiation," M. Grandolfo, S.M. Michaelson and A. Rindi, eds. Plenum, NY (1983).

7. W.R. Adey, Molecular aspects of cell membranes as substrates for interaction with electromagnetic fields, *in*: "Synergetics of the Brain," H. Haken and A. Mandell, eds. Springer, Berlin (1983).

8. R. Dixey and G. Rein, ^{3}H-nonadrenaline release potentiated in a clonal nerve cell line by low-intensity pulsed magnetic fields. *Nature* *296*:253 (1982).

9. A.F. Lawrence and W.R. Adey, Nonlinear wave mechanisms in interactions between excitable tissue and electromagnetic fields. *Neurol. Res.* *4*:115 (1982).

10. L.K. Kaczmarek and W.R. Adey, Weak electric gradients change ionic and transmitter fluxes in cortex, *Brain Res.* *66*:537 (1974).

11. S.M. Bawin and W.R. Adey, Sensitivity of calcium binding in cerebral tissue to weak environmental electric fields oscillating at low frequency, *Proc. Natl. Acad. Sci. USA* *73*:1999 (1976).

12. S.M. Bawin, A.R. Sheppard and W.R. Adey, Possible mechanisms of weak electromagnetic field coupling in brain tissue, *Bioelectrochem. Bioenergetics* *5*:67 (1978).

13. S.M. Bawin, W.R. Adey and I.M. Sabbot, Ionic factors in release of $^{45}Ca^{2+}$ from chick cerebral tissue by electromagnetic fields, *Proc. Natl. Acad. Sci. USA* *75*:6314 (1978).

14. W.R. Adey, S.M. Bawin, and A.F. Lawrence, Effects of weak amplitude-modulated microwave fields on calcium efflux from awake cat cerebral cortex, Bioelectromagnetics 3:295 (1982).

15. S. Lin-Liu and W.R. Adey, Low frequency amplitude modulated microwave fields change calcium efflux rates from synaptosomes, Bioelectromagnetics 3:309 (1982).

16. R.A. Luben, C.D. Cain, M.C.-Y. Chen, D.M. Rosen and W.R. Adey, Effects of electromagnetic stimuli on bone and bone cells in vitro: inhibition of responses to parathyroid hormone by low-energy, low-frequency fields, Proc. Nat. Acad. Sci. USA 79:4180 (1982).

17. C.V. Byus, R.L. Lundak, R. Fletcher and W.R. Adey, Exposure to a modulated microwave field results in a decrease in protein kinase activity in human lymphocytes, Proc. 5th Annual Meeting, Bioelectromagnetics Soc., p. 2 (abstract) (1983).

18. L. Bass and W.J. Moore, A model of nervous excitation based on the Wien dissociation effect, in: "Structural Chemistry and Molecular Biology," A. Rich and C.M. Davidson, eds. Freeman, San Francisco (1968).

19. W.R. Adey, Frequency and power windowing in tissue interactions with weak electromagnetic fields, Proc. IEEE 68:119 (1980).

20. J.L. Schwartz, J. Delorme, G.A.R. Mealing, Effects of low-frequency amplitude modulated radio frequency waves on the calcium efflux of the heart, Biophys. J. 41:295a (1983) (abstract).

21. D.B. Lyle, P. Schechter, W.R. Adey and R.L. Lundak, Suppression of T-lymphocyte cytotoxicity following exposure to sinusoidally amplitude-modulated fields, Bioelectromagnetics 4(3):(1983).

22. J.R. Lymangrover, E. Keku and Y.J. Seto, 60 Hz electric field alters the steroidogenic response of rat adrenal tissue, in vitro, Life Sciences 32:691 (1982).

23. S.M. Bawin, L.K. kaczmarek and W.R. Adey, Effects of modulated VHF fields on the central nervous system, Ann. NY Acad. Sci. 247:74 (1975).

24. C.S. Blackman, J.A. Elder, G.M. Weil, S.G. Bennane, D.C. Eichinger and D.E. House, Induction of calcium ion efflux from brain tissue by radio frequency radiation; effects of modulation frequency and field strength, Radio Sci. 14:93 (1979).

25. C.S. Blackman, S.G. Bennane, D.E. House, W.T. Joines, and R.J. Spiegel, Further developments in ELF-induced changes in calcium-ion efflux from brain tissue. Proc. Bioelectromagnetics Soc., 4th Annual Meeting, p. 30 (abstract) (1982).

26. A. Katchalsky, Polyelectrolytes and their biological interactions, in: "Connective Tissue: Intercellular Macromolecules," NY Heart Assocn., ed. NY Heart Assocn., New York (1964).

27. C.S. Blackman, S.G. Bennane and L.S. Kinney, Effects of ELF radiation on calcium ion efflux from brain tissue in vitro. Proc. Bioelectromagnetics Soc., 3rd Annual Meeting, p. 47 (abstract) (1981).

28. W.B. Jolley, D.B. Hinshaw, K. Knierim, and D.B. Hinshaw, Electromagnetic field effects on calcium efflux and insulin secretion in isolated rabbit islets of Langerhans, Bioelectromagnetics 4:103 (1983).

29. C.P. Taylor and F.E. Dudek, Synchronous neural after discharges in rat hippocampal slices without active chemical synapses, Science 218:810 (1982).

30. J.G.R. Jefferys and H.L. Haas, Synchronized bursting of CA_1. hippocampal pyramidal cells in the absence of synaptic transmission, Nature 300:448 (1982).

31. H. Frohlich, Coherent electric vibrations in biological systems and the cancer problem, IEEE Trans. Microwave Theory Tech. 26:613 (1978).

32. W.R. Adey and S.M. Bawin, Brain interactions with weak electric and magnetic fields, Neurosci. Res. Program Bull. 15:1 (1977).

33. I.T. Grodsky, Neuronal membrane: a physical synthesis, Math. Biosci. 28:191 (1976).

34. R.R. Birge, Photophysics of light transduction in rhodopsin and bacteriorhodopsin, Ann. Rev. Biophys. Bioeng. 10:315 (1981).

35. A.S. Davydov, Solitons in molecular systems, Physica Scripta 20:387 (1979).

36. J.M. Hyman, D.W. McLaughlin and A.C. Scott, On Davydov's alpha-helix soliton, Physica D 30:23 (1981).

37. A.C. Scott, The laser-Raman spectrum of a Davydov soliton, *Physics Letters 86A*:60 (1981).

USE OF BONE CELL HORMONE RESPONSE SYSTEMS TO INVESTIGATE BIOELECTROMAGNETIC EFFECTS ON MEMBRANES *IN VITRO*

Richard A. Luben and Christopher D. Cain

Division of Biomedical Sciences and
Department of Biochemistry,
University of California
Riverside, CA 92521

INTRODUCTION

In a number of clinical studies (1-3) the usefulness of induced electromagnetic fields in stimulating healing of chronic fracture non-unions has been demonstrated. These fields stimulate a number of activities including the synthesis of new matrix by osteoblasts (4), transcription of genetic information by bone cells (5) and other processes associated with fracture healing. In addition, DC voltage applied to bones by invasive techniques has been reported to stimulate both localized osteolysis (in areas near the anode), and localized osteogenesis (in areas near the cathode), presumably by alteration of the activity of endogenous bone forming and/or bone resorbing cells (6). Externally induced electromagnetic fields used for treatment of fractures produce tissue current densities of less than 1 microampere/cm^2 and associated electric gradients of about 1 mV/cm (2). As yet there is limited understanding of the possible mechanisms by which fields of this low magnitude (relative to tissue transmembrane potentials and cellular impedances) could bring about such significant changes in the activity of bone cells *in situ*.

In previous studies (4) this laboratory has examined the activities of isolated bones and of osteoblast-like cells in culture during treatment with induced pulsing electromagnetic fields identical to those used for treatment of fractures. Our initial findings were that these fields produced substantial inhibition of the response of bone cells to parathyroid hormone (PTH), a hormone which generally induces osteolytic effects in bone. Moreover, the data indicated that the actions of fields were localized to the plasma membrane of the cells, interfering with the activation of adenylate cyclase by PTH bound to its cell-surface receptor. Based on these findings, we have developed a hormonal hypothesis for the mechanism of action of electromagnetic fields in bone. This hypothesis suggests that in vivo activation of fracture healing may be brought about by induction of localized insensitivity of bone cells to the actions of circulating PTH, thus leading to a net increase in bone formation in the treated area. In the current studies, we have further examined this proposed mechanism by studying the effects of fields on individual molecular steps of hormone action in bone cells. The current results strengthen the indication that electromagnetic fields act by changing the properties of the cell membrane, thus changing the ability of cells to respond to regulatory factors in their environment. The data also suggest that the effects of fields on membranes are complex, with marked dependencies on duration of exposure and time of measurement of the observed phenomena. These findings further illustrate the concept that fields of very low energy can exert profound effects on metabolic activities by interacting with the transduction of regulatory signals across the cell membrane.

MATERIALS AND METHODS

To furnish intact isolated bone samples for experiments, cranial bones were dissected from 3-day old mice and cultured as previously described (7). In other experiments, monolayer cultures of osteoblasts were employed using the MMB-1 osteoblast cell line isolated by this laboratory (8,9) from primary cultures of mouse bone cells. Cells were cultured in Eagle's MEM medium containing 10% fetal bovine serum (Gibco), in an atmosphere of 5% CO^2 in air at 37°C. To furnish replicates for experiments, cells were subcultured to multiwell plates (Falcon) either at 20,000 cells per 25-mm well or at 10,000 cells per 16-mm well. Cells were replaced in the incubator and were cultured in control medium for at least 18 hours before treatment. Dishes to be treated with electromagnetic fields were placed in the central space of a 15-cm diameter Helmholtz coil attached to a Bi-Osteogen (tm) clinical field generator (furnished by Electro-Biology, Inc., Fairfield, New Jersey). The coil was placed in a tissue culture incubator, while the pulse generator unit remained outside the incubator at normal room temperature and humidity. Control cultures were placed in the same incubator as the treated cultures, but were isolated from the field by layers of styrofoam and stainless steel. These isolation conditions produced at least 99.99% attenuation of the field as measured by induction probes placed at various locations in the incubator (4).

For measurement of cyclic AMP production, after specified times in the fields the culture medium was replaced with fresh medium containing 0.01 mM theophylline. Cells or bones were then incubated for 30 min to allow temperature and gas equilibration before hormone treatment. PTH was then added for 5 min, after which the cells or bones were rapidly killed by exposure to a commercial microwave oven (7). Assays of cyclic AMP accumulation or adenyl cyclase activity were performed as described previously (4,7). The PTH used in these studies was bovine PTH 1-34 synthetic peptide (Bachem). The electromagnetic field used for these experiments has been described previously (4,10). The waveform of the field designated as SPP ("single pulse, patient") is generated by a single pulse of positive current applied to the induction coil for 325 usec at 18.9 volts, followed by a broad negative "tail" over 5 msec. The pulse is repeated at a frequency of 72 Hz. This field has been used clinically to improve healing of chronic non-unions and pseudoarthroses (1). At the level of tissues or cells exposed both in vivo and in vitro, the field produces an electrical gradient of about 1 mV/cm and a current density of approximately 1 uA/cm^2 (4). The amplitude, waveform and frequency of the field was confirmed periodically using a miniature induction coil probe (10). In all experiments described here, the cultures were exposed to the field continuously for the indicated times prior to testing for hormone responses. In experiments where exposure was terminated prior to measuring cAMP formation, the field was turned off at the generator without disturbing the cultures; incubation was continued for the specified periods before addition of hormone and cAMP measurement as described above. All experiments were carried out on triplicate or quadruplicate cultures, and significance of the results was analyzed by Dunnett's t-test.

For studies of the binding of PTH to bones, the PTH peptide was labeled with ^{125}I by a chloramine-T procedure and purified by HPLC as described previously (11). Bones were incubated at 4°C in the presence of radioactive PTH plus increasing quantities of unlabeled PTH. All incubations were carried out for 3 hours to insure equilibrium binding, after which the bones were washed three times with cold saline and counted for bound ^{125}I. Scatchard plots were constructed to measure the binding affinity and number of sites for PTH.

RESULTS

When isolated mouse bones were cultured in the presence of the described electromagnetic field, PTH-stimulated accumulation of cAMP was markedly reduced (Table 1), as previously reported (4). This inhibition was persistent, as demonstrated by the fact that it could be observed for at least three hours after turning off the field. It should be noted that in all cases exposure of bones to the hormone was for a standardized period of five minutes, regardless of the time of exposure to field or the time after turning off the field. In other experiments (not shown) we had ascertained that the time course (as opposed to the magnitude) of adenyl cyclase activation by PTH was not altered by exposure to the field and that five minutes remained the time of maximum activation, as shown previously (7).

TABLE 1
Cyclic AMP Responses of Mouse Bone to PTH Treatment
During and After Field Exposure

PTH Dose	cAMP, pmol/bone In Field	 Out of Field
None	0.3 ± 0.1	0.4 ± 0.2
10 ng/ml	0.4 ± 0.1	0.4 ± 0.1
30 ng/ml	1.3 ± 0.4*	0.7 ± 0.2
100 ng/ml	3.1 ± 0.5*	0.8 ± 0.3

Cranial bones from four day old mice were cultured for 48 hours in the presence of the electromagnetic field described in Methods. At the end of 48 hours, some bones were removed from the field for one hour ("Out of Field") prior to treatment with the indicated doses of PTH. The remainder of the bones were treated with PTH while still in the field ("In Field"). Values expressed are the means ± standard errors for quadruplicate samples. Asterisks denote significant difference ($P < .05$) between treatment groups at a given dose of hormone. The values for bones cultured in the absence of any field exposure (not shown) were not significantly different at any dose from the bones designated as "In Field".

To our considerable surprise, however, the expected field-mediated inhibition of cAMP accumulation was not observed when PTH treatment was carried out while bones were still exposed to the field, or immediately after removing bones from the field. This was in contrast to our previous experiments in this system (4), in which PTH treatments had always been carried out 45-60 minutes after removal of samples from the field. Further examination of this phenomenon revealed that no inhibition of PTH response could be observed earlier than about 30 minutes subsequent to turning off the field (Table 2). The observed inhibition of PTH responsiveness reached its maximum within one hour of continued culture of bones in the absence of the field, and persisted for approximately four hours, after which normal levels of responsiveness were regained. Although unexpected in this system, other examples of what might be termed "time-windowing" in the effects of electromagnetic fields have been described in other systems (see Discussion).

TABLE 2
Time Course of Inhibition of PTH Response
After Termination of Field Exposure

Time Out of Field	cAMP, pmol/bone
0	3.4 ± 0.6
1 hr	1.8 ± 0.4*
2 hr	1.9 ± 0.3*
3 hr	1.7 ± 0.4*
4 hr	3.1 ± 0.4
No Field	3.6 ± 0.5

Cranial bones from four day old mice were cultured for 48 hours in the presence of the electromagnetic field described in Methods. At the end of 48 hours, the field was turned off and culture was continued for the indicated times before treatment for 5 min with 100 ng/ml PTH. For the point designated zero time out of field, treatment with PTH was for the first five minutes after turning off the field. Values expressed are the means ± standard errors for quadruplicate samples. Asterisks denote significant difference ($P < .05$) from the "No Field" group. The values for bones not treated with PTH were not affected by this time course and were consistently 0.5 ± 0.3 pmol/bone for all time points.

It was of interest to exmine more closely the possible mechanisms by which such complex alterations in membrane response could be brought about. We had shown in previous studies (4) that the total catalytic activity of adenyl cyclase in the membranes, as measured by activation with fluoride, was not altered by culture of cells in electromagnetic fields. A reasonable alternative hypothesis was that the field might alter the number, affinity or availability of cell surface hormone receptors rather than the cyclase enzyme itself. Consequently, we examined the binding to bone cells of an ^{125}I-labeled preparation of PTH which has been shown (11) to have identical receptor-binding and biological activities to the authentic hormone. These studies (Table 3) indicated that the field failed to change significantly either the number or affinity of receptors for the hormone under the experimental conditions employed.

TABLE 3

Binding Constants of ^{125}I-PTH to Mouse Bone

Treatment	$^{*}K_D$	$^{**}B_{max}$	$^{***}R$
No Field	1.45×10^{-12}	5.88×10^{-15}	0.82
Field (14 hr)	2.07×10^{-12}	9.90×10^{-15}	0.78

*Dissociation constant, molar.
**Maximum binding intercept, moles hormone/mg dry bone.
***Correlation coefficient.

Mouse cranial bones were cultured for 14 hours in the presence or absence of the electromagnetic field. The bones were then removed from the incubator and placed in culture dishes which had been treated with poly-L-lysine (1 mg/ml) to minimize nonspecific adsorption of hormone. To each dish was added 300,000 cpm of ^{125}I-PTH (nle^{8},nle^{18},tyr^{34} derivative of bovine PTH) and increasing concentrations of unlabeled hormone, dissolved in culture medium with 2 mg/ml bovine serum albumin. Incubation was carried out at 4°C (not in the field) for 3 hr. After washing with ice-cold saline, the bones were dried, weighed and the radioactivity counted. Scatchard analysis was performed using a computer program for least-squares fitting of the data, plotted as bound/free vs. bound hormone (expressed as fmol/mg dry bone).

Although these data do not eliminate the possibility that there might be a more subtle change in the receptor under field treatment (e.g., a steric change in some portion of the receptor not directly involved in hormone binding), they suggest that direct interference with hormone binding is not the major mechanism of action of the field.

The data described above argued at least tentatively against a primary role in field actions of two components of the cell membrane hormone-cyclase complex, namely the hormone receptor protein and the adenylate cyclase catalytic protein. However, a third component also is known to participate in the actions of PTH on adenyl cyclase: the nucleotide binding regulatory protein, commonly known as the N-protein (12). To determine whether the activity of this protein was altered by the field, we examined the response of adenyl cyclase to cholera toxin, an agent which activates the cyclase molecule by covalently modifying the N-protein (12). The results indicated (Table 4) that the field produced a significant increase in the ability of cholera toxin to activate cyclase in intact bone cells at submaximal doses, although there was no significant effect on the maximal level of activation attainable at higher doses.

TABLE 4
Effect of Electromagnetic Field on cAMP Response of MMB-1 Bone Cells to Cholera Toxin

Toxin Dose	cAMP, pmol/bone No Field	 Field (48 hr)
None	0.6 ± 0.4	1.3 ± 0.4
10 ng/ml	1.0 ± 0.5	3.7 ± 0.8*
100 ng/ml	1.8 ± 0.3	7.8 ± 1.0*
1 ug/ml	6.8 ± 1.2	8.3 ± 1.3*

MMB-1 cells were cultured for 48 hours in the presence or absence of the electromagnetic field described in Methods. At the end of 48 hours, the field was turned off and all cells were treated for 20 min with 5 mM theophylline to inhibit endogenous phosphodiesterase, and then for 45 minutes with the indicated doses of cholera toxin prior to assay of cyclic AMP accumulation (see Mthods). Values expressed are the means ± standard errors for quadruplicate samples. Asterisks denote significant difference ($P < .05$) between treatment groups at a given dose of toxin.

These effects were not observed in rat hepatoma cells exposed to the field, although a substantial level of cyclase activation by cholera toxin was observed (data not shown). It also should be noted that the effects of the field on cholera toxin sensitivity of intact bone cells were persistent; the experiments described in Table 4 were carried out after at least one hour of culture in the absence of field exposure. Interestingly, however, the effects observed on intact cells were not demonstrated in isolated bone cell membranes treated with cholera toxin active subunits (data not shown), suggesting that some dynamic property of the intact living cell membrane might be crucial in the ability of fields to influence cell metabolism.

DISCUSSION

PTH causes markedly increased levels of cAMP in isolated bone and bone cells treated in culture, as demonstrated by numerous studies (13). The general direction of metabolic effects of PTH in acute doses (as employed in these studies) is toward osteolysis, although there appear to be long-term osteogenic effects of low doses of PTH as well. Thus, localized interference with the short-term osteolytic effects of PTH *in vivo* might be expected to result in a localized increase in bone formation. This in turn would be consistent with the observed clinical effects of the electromagnetic fields employed in this series of studies. In previous studies we showed that, when culture of bone and bone cells was carried out in the presence of electromagnetic fields, the PTH-stimulated accumulation of cAMP in both intact isolated cranial bones and MMB-1 osteoblasts was markedly reduced. This inhibition took the form of a significant decrease in the sensitivity of the cells to hormone; the concentration of PTH required for half-maximal activation of adenyl cyclase was increased by 50-100 fold. The observed inhibition of adenyl cyclase activation was accompanied by parallel inhibition of subsequent intracellular actions of PTH. In the same studies, we found that the observed inhibition of PTH-dependent adenylate cyclase activation was not caused by loss of cyclase catalytic subunit from the membranes. Moreover, the inhibition of cyclase response by electromagnetic fields extended to other agents (e.g., osteoclast activating factor) which require activation of membrane adenyl cyclase for their effects. On the other hand, agents which do not appear to act via membrane cyclase-receptor mechanisms (e.g., 1,25-dihydroxyvitamin D) were not affected.

In the current studies, we have extended the above findings by examining in closer detail the individual steps of the membrane receptor-cyclase system. The data reported here indicate that fields probably do not directly interfere with the binding of PTH to its membrane receptor either in terms of affinity or total number of receptors (although we cannot rule out at this point a change in the receptor which would not be observed as a change in hormone-binding properties). Moreover, the adenyl cyclase catalytic subunit itself appears by all measures so far investigated to be normally functional in cells treated with fields. Our preliminary evidence does, however, implicate the regulatory N-protein as a possible mediator of the actions of the field. This protein appears to regulate the formation of a ternary complex between three normally non-associated proteins in the cell membrane: the hormone receptor (sterically activated by hormone binding); the N-protein itself; and the catalytic adenylate cyclase molecule, which becomes active as a result of association with the N-protein. The activity of the N-protein is apparently under multiple influences (12), including its ability to bind GTP and other nucleotides. Moreover, there is evidence that more than one N-protein may reside in the membrane of a given cell type, further complicating the regulatory picture (14).

It is not as yet clear from our observations what the mechanism of field influences on N-protein function may be. A direct effect by field-induced voltages or magnetic fluxes on protein structure or metabolism would seem to be rather unlikely due to the extremely small current and voltage involved (10). Yet another important consideration is that effects of the field on N-protein function might not necessarily be on the protein itself. Changes in the surface ionic properties of the membrane or changes in membrane fluidity might change the association properties of the N-protein for other components of the receptor-cyclase complex. A considerable amount of investigation will be required to examine all these possibilities. It is clear, however, that whatever the changes in membrane function might be, they are persistent, dynamic and specific to individual cell types.

We cannot as yet explain the significance of our observations that bone cell membrane function remains apparently intact _during_ exposure to the field, and that field effects only become apparent _after_ turning off the field. One possibility is that multiple cellular adaptive functions are involved in maintaining homeostatic control during exposure to the field; and that the observed transient loss of sensitivity to hormone is a function of a kinetic disequilibium which occurs during readjustment to the absence of the field. It is worthy of note in this regard that clinical effectiveness of the fields appears to be enhanced markedly by intermittent, rather than continuous exposure of fracture sites to the field (1). There is ample precedent for persistent changes in membrane properties as a result of exposure to low-energy electromagnetic fields, most probably due to long-range rearrangements of the ionic layers at (or in the solvent immediately adjacent to) the membrane surface (15).

It is our belief that the current system offers a useful model for future examination of the molecular basis of field effects at a number of levels, ranging from whole-organ changes in metabolism to precise molecular details of hormone-receptor interactions and other processes. As such, bone and bone cells _in vitro_ may furnish important information both on the basic electrochemical processes and on the clinically relevant phenomena associated with bioelectromagnetic phenomena.

ACKNOWLEDGEMENTS

We gratefully acknowledge the technical assistance of Nora Vasquez, and the advice, assistance and support of W.R. Adey. These studies were carried out under a contract from the DOE (No. DE AI01-79ET29078) and with the aid of generous gifts of funds and equipment from Electro-Biology, Inc.

REFERENCES

1. Bassett, C.A.L., Pilla, A.A. and Pawluk, R.J. A non-operative salvage of surgically-resistant pseudarthroses and non-unions by pulsing electromagnetic fields. Clin. Orthop. Rel. Res. 124:128-142 (1977).
2. Brighton, C.T., Black, J., Friedenberg, Z.B., Esterhai, J.L., Day, L.J. and Connoly, J.F. A multicenter study of the treatment of non-union with constant direct current. J. Bone and Joint Surg. 63-A:2-12 (1981).
3. Becker, R.O. Electrical osteogenesis--pro and con. Calc. Tiss. Res. 26:93-97 (1978).
4. Luben, R.A., Cain, C.D., Chen, M.C.Y., Rosen, D.M., and Adey, W.R. Effects of electromagnetic stimuli on bone and bone cells in vitro: Inhibition of responses to parathyroid hormone by low-energy low-frequency fields. Proc. Nat. Acad. Sci. (USA) 79:4180-4184 (1982).
5. Bassett, C.A.L. Pulsing electromagnetic fields: A new method to modify cell behavior in calcified and noncalcified tissues. Calc. Tiss. Internat. 34:1-8 (1982).
6. Bassett, C.A.L. The biological significance of piezoelectricity. Calc. Tiss. Res. 1:251-272 (1968).
7. Luben, R.A., Chen, M.C.Y., Rosen, D.M. and Mohler, M.A. Effects of osteoclast activating factor from human lymphocytes on cyclic AMP concentrations in isolated mouse bone and bone cells. Calc. Tiss. International 28:23-32 (1979).
8. Rosen, D.M., and Luben, R.A. Multiple hormonal mechanisms for the control of collagen synthesis in an osteoblast-like cell line, MMB-1. Endocrinology 112:992-999 (1983).
9. Walters, M.R., Rosen, D.M., Norman, A.W., and Luben, R.A. 1,25-Dihydroxyvitamin D receptors in an established bone cell line: Correlation with biochemical responses. J. Biol. Chem. 257:7481-7484 (1982).
10. Pilla, A.A. in "Bioelectrochemistry," ed. Keyzer, H. and Gutman, F. (Plenum Publishing Corp., New York), pp. 353-396 (1980).

11. Rizzoli, R.E., Somerman, M., Murray, T.M., and Aurbach, G.D. Binding of radioiodinated parathyroid hormone to cloned bone cells. Endocrinology 113:1832-1838 (1983).
12. Ross, E.M. and Gilman, A.G. Reconstitution of catecholamine-sensitive adenylate cyclase activity: Interaction of solubilized components with receptor-replete membranes. Proc. Nat. Acad. Sci. (USA) 74:3715-3719 (1977).
13. Peck, W.A. Cyclic AMP as a second messenger in the skeletal actions of parathyroid hormone: A decade-old hypothesis. Calc. Tiss. International 29:1-4 (1979).
14. Kirchick, H.J., and Birnbaumer, L. Effects of estradiol treatment on rabbit luteal adenylyl cyclase: Loss of luteinizing hormone receptors and attenuation of the regulatory N component activity. Endocrinology 113:1629-1637 (1983).
15. Adey, W.R. Tissue interactions with nonionizing electromagnetic fields. Physiol. Revs. 61:435-514 (1981).

STRONG INTERACTIONS OF RADIOFREQUENCY FIELDS WITH NUCLEIC ACIDS

Mays L. Swicord and Przemyslaw Czerski

Electromagnetic Radiation Branch, Div. of Life Sciences
National Center for Devices and Radiological Health
Rockville, Maryland 20857

Glenn S. Edwards*

Chemical Physics Program, University of Maryland

Christopher C. Davis

Electrical Engineering Dept., University of Maryland

INTRODUCTION

We first present evidence of genetic effects observed in various organisms exposed to radiofrequency (RF) radiation. Then we will change the subject and discuss findings on the absorption characteristics of DNA in solution. The relation between these subjects, RF genetic effects, and the strong interactions of RF with the DNA molecule is speculative. It is indeed an area ripe for theoretical investigation. We conclude this presentation with remarks suggesting that non-linear processes may be a key to understanding the relation between rapidly thermalizing models in DNA and genetic effects.

REVIEW OF GENETIC EFFECTS

Genetic effects resulting from point mutations are unlikely due to the small photon energies (i.e. 4×10^{-4} ev at 10 GHz) in the radiofrequency range. Most investigators have not observed such phenomena, and in fact, low photon energy is sometimes used as an argument for the non-existence of any radiofrequency genetic effects. However, RF induced genetic changes other than point mutations have been observed. These include nonheritable effects, such as effects on gene expression, i.e., alterations in cellular functions, which are under genetic control. Heritable changes, i.e., RF induced alterations in the number and/or structure of chromosomes in somatic animal and plant cells were first reported by Heller and Teixeira-Pinto,[1] almost a quarter of a century ago. These findings were confirmed and expanded by many authors, e.g., Heller,[2] Kapustin et al.,[3] Janes et al.,[4] Yao and Jiles,[5] Stodolnik-Baranska,[6] Yao,[7], [8] Cieciura et al.,[9] and Chen et al.[10] (see Leach[11] for review). Aberrations which have been observed include changes in the number of chromosomes (aneuploidy), formation of dicentrics, gaps or chromosome breaks, acentric fragments, and translocations. The majority of this work was done at frequencies between 2 and 10 GHz.

---

*Contributions to this paper represent work to be submitted in partial fulfillment of the requirements for the Ph.D. degree.

All of the above mentioned aberrations are also observed when the biological system is exposed to toxic chemicals or ionizing radiation (Hirschhorn and Cohen[12]). It has been suggested that these observed chromosomal aberrations are due to heating of the system by RF energy. However, most thermal experimental work indicates that heat cell killing occurs before chromosomal aberrations can be induced by heat (Dewey et al.,[13,14] Westra and Dewey[15]). Heat does, however, induce point mutations. (Mitler,[16] Smith and Corwin[6]).

There is one observed aberration that is unique to RF energy exposure - despiralization - the uncoiling of chromosomes, which may suggest the breakage of histone bonds. This despiralization due to

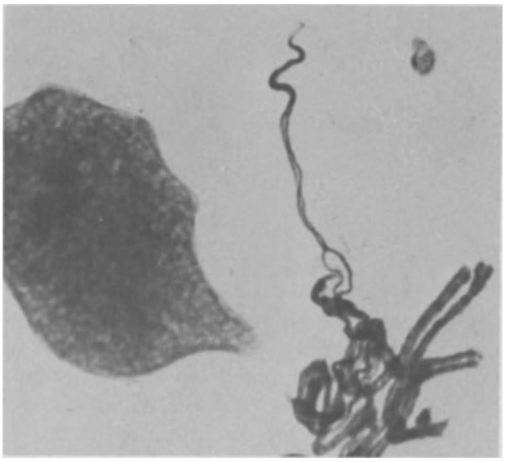


Figure 1 - Despiralization of chromosomes in rat kangaroo cells exposed to 2.45 GHz radiation (Yao and Jiles[5]).

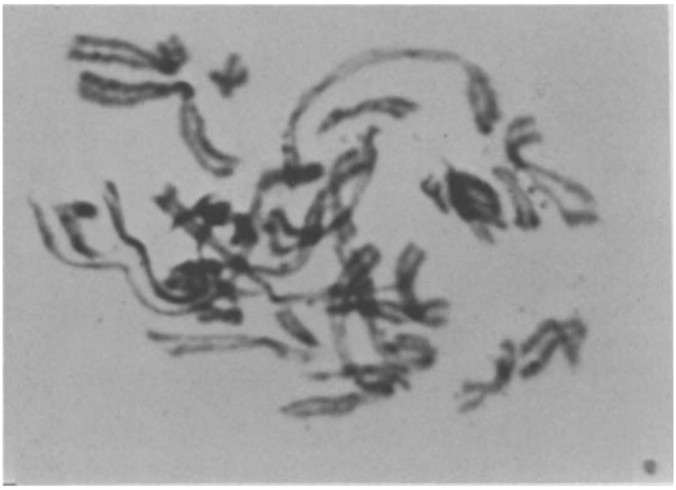


Figure 2 - Despiralization of chromosomes in human lymphocytes exposed to 2.95 GHz radiation (Stodolnik and Baranska[6]).

RF exposure was observed independently by different investigators. Yao and Jiles[5] observed despiralization in rat kangaroo cells exposed to 2.45 GHz radiation (Fig 1) and Stodolnik-Baranska[6] observed the same phenomena in human lymphocytes exposed to 2.95 GHz radiation (Fig 2). The uncoiling phenomenon, which has been described only following microwave exposure, could possibly suggest the excitation of modes not apparently coupled to the thermal bath, an unusual interaction with the forces holding together the DNA and mammalian chromosome material.

Much research has been done on the response of male meiotic chromosomes to microwave exposure. Effects on spermatogenesis as well as aberrations such as translocations have been observed by Manikowska-Czerska et al.[18,19,20,21] by direct examination of sperm cells from mice exposed to 915 MHz, 2.45 and 9.4 GHz. The genetic significance of these findings has been demonstrated by effects on offspring of males exposed at 915 MHz.[12] Dominant lethal tests were performed by intrauterine examination of offspring resulting from the breeding of exposed male mice with unexposed females. Observations demonstrated either an increase in early and late fetal death resulting in a decrease of normal live offspring or the sterility of exposed males. Similar experiments were performed with the same results in at least two other laboratories (Varma and Traboulay,[23] Varma et al.,[24] Gould et al[25]). Negative findings were reported by Berman et al[26] with rats as the experimental subjects. In addition, certain types of translocations induced in fathers were recovered in first generation male offspring, indicating heritable translocations. (Manikowska-Czerska[18] and personal communication).

Research has characterized spermatogenesis in mice as a funtion of time development[27,28]. Figure 3 depicts effects of 2.45 GHz ratiation on the number of divisions of spermatogonia and speratocytes I and II as reported by Manikowska-Czerska et al.[19,20]. The independent variable is the rate of whole body averaged absorbed energy (specific absorption rate or SAR). Dosimetric studies by HO and Czerski (unpublished) indicate that the averaged absorbed energy rate in the testes was higher approximately by a factor of two than the whole body average for the same experimental situations. The rate of average whole body absorbed energy ranged from 0.05 W/kg to 20 W/kg in these studies. Core temperature measurements showed increases of 0.1 to 0.2°C at 20 W/kg. However, measurements of temperature in the testes, which is a principal organ of thermoregulation and heat dissipation in the mouse, showed no changes at the highest level (20 W/kg) of exposure. Exposures were for 30 minutes a day, 6 days a week for 2 weeks. The ambient temperature was 24°C, the relative humidity was 30% and volume of air flow through the waveguide was 15 liters per minute. Examination of cells occurred 2 days following the final day of exposure. Note the immediate decrease in the count of the spermatogonial metaphases (MG) and the accompanying decrease of spermatocytes I and II. These results become significant at 0.5 W/kg. Metaphases counts in all stages increase until there is no observable effect at 5 W/kg. The increase continues with increasing absorbed power. At 20 W/kg, a decrease in metaphase II counts occurs. This can be explained by accumulation of chromosome aberrations at metaphase I and the inability of cells to progress from metaphase I to metaphase II. Note that an experiment performed at the single level of 5 W/kg would find no significant effect. One can only speculate at this point as to the reason for the reversal of the effect. However, such reversals are not uncommon in responses of proliferating and differentiating cell populations or genetic responses to various insults.

For example, recently Aarholt et al.[29] reported a reversal in the relative rate of β-Galactosidase synthesis of E. Coli when exposed to increasing magnetic field strengths. The fields were square wave modulated at 50 Hz. There was a decrease in the rate of synthesis at 0.27 mT and an increase at 0.58 mT and both were sharply dependent on field strength. This of course is an example of an effect on gene expression to be discussed further, and not an example of a change in genetic information content (i.e., a heritable change). However, both types of genetic effects may possibly take place if the system is exposed during the time the chromosome is in a delicate functional state, such as meiosis. The translocations observed by Manikowska-Czerska

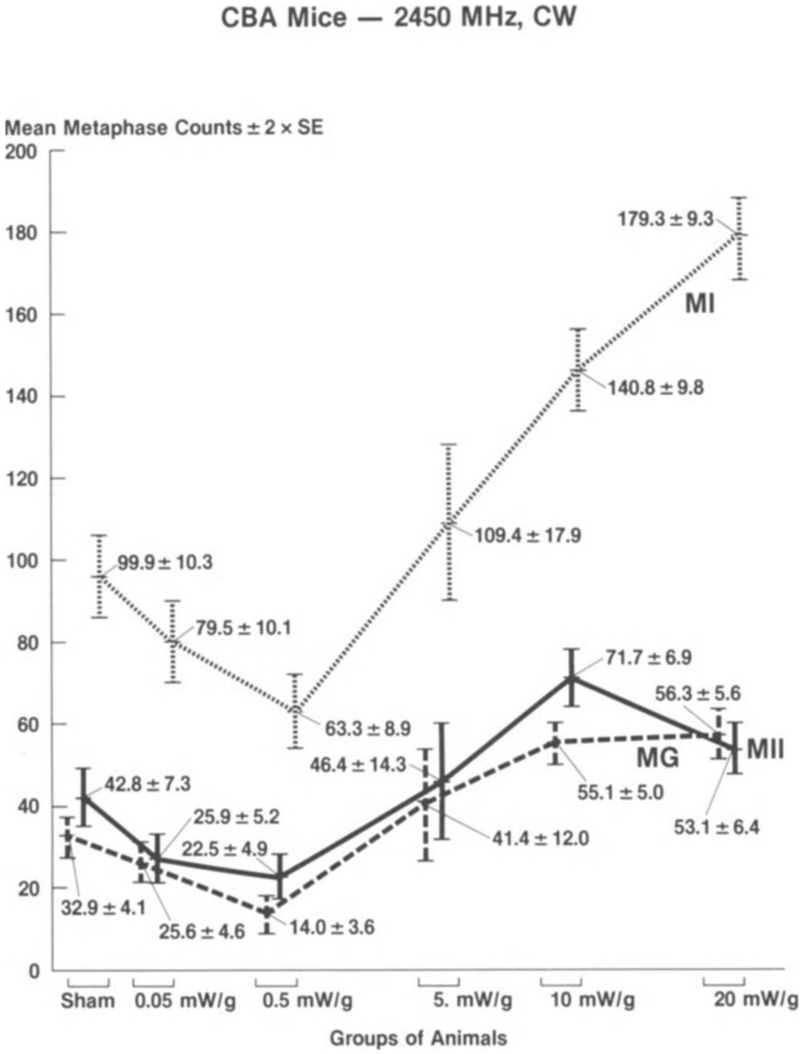


Figure 3 - The effects observed in mice of 2.45 GHz radiation on the number of divisions of spermatogonia and spermatocytes I and II.

et al.[19,20] from exposure to 2.45 GHz radiation during spermatogenesis.

The expected occurrence of spontaneous translocation is approximately 3 to 9 in 10,000 animals.[30] Data of Manikowska-Czerska et al. in Table I indicate an occurrence in 7% of cells in 6 animals examined per group, each animal having at least one cell with translocations. The number of translocations does increase in a non-linear manner with increases in the intensity of exposure.

Table 1

Cell with Translocations in Shams and in Experimental Animals

Dose-Rate	Total Cells Analysed	Translocations	
		Total	Percent
Sham	248	0	0
0.05 mW/g	111	8	7.2
0.5 mW/g	116	8	6.9
5 mW/g	214	9	4.2
10 mW/g	105	13	12.4
20 mW/g	122	15	12.3

Inactivation of repressor molecules with RF radiation and the resulting activation of gene expression have been indicated by the work of several investigators. Frequency dependent activation of colicins genesis was reported by Smolyanskaya et al.[31] and discussed by Keilmann at this conference. Webb[32] has reported λ phage induction by exposure to radiation in the 70 GHz frequency range. The work of Aarholt et al.[29], quoted earlier, should be included, since transcriptions of β-Galactosidase is controlled by a repressor protein. In addition, puffing of giant chromosomes has been observed by Goodman et al.[23] at ELF frequencies, and at millimeter wave frequencies by Kremer et al.[34] Cell didifferentiation resulting from exposure to ELF has been reported by Chiabrera et al.[35]

One should not conclude from this brief review that radio-frequency waves in general produce genetic effects. This is not the case. These reported effects are frequency dependent, amplitude dependent, and sometimes modulation-dependent.

For completeness, one should recognize that negative effects also have been reported. By-and-large, these negative findings seem to be associated with the search for RF induced point mutations. Sister chromatid exchanges, which have been suggested as an indicator of the probability of the occurrence of point mutations were investigated by McRee et al.[36] with negative findings. Baranski et al.[37], Blackman et al.[38], Dardalhon et al.[39], and Dutta et al.[40] have investigated mutations in unicellular organisms with negative findings. Mittler[41] and Pay et al.[42] were unable to find mutagenic effects in Drosophila. These results suggest that point mutations do not occur due to radiofrequency radiation, which seems to be a reasonable conclusion. However, it must also be pointed out that no systematic examination for the induction of point mutations over a range of RF doses has been made.

DNA ABSORPTION

We now change the subject to a discussion of the microwave absorption properties of the DNA molecule. We have extracted purified DNA from E. coli using classical procedures (discussed in reference 43) and measured its absorption properties (primarily in the X-band region) using two distinctly different measurement techniques and two variations of the second. The first technique has been termed by its originator, Chris Davis, phase fluctuation optical heterodyne spectroscopy or PFLOH spectroscopy.[44] The method was originally developed for trace gas detection. We have found the method applicable to the study of microwave absorption by aqueous solutions.[45,46] Although we applied the method in the X-band, 8 to 12 GHz, we feel it will be most useful at higher frequencies.

The schematic arrangement is shown in Figure 4. A Mach-Zehnder configuration is chosen as the basic interferometer, although others can be used. The light from a single-frequency He-Ne laser is divided by a beam splitter (B1) into a signal and reference beam and recombined into a collinear beam after traversing approximately equal path lengths. The combined signal is directed onto a photodetector. The signal beam passes through the liquid sample (contained in 1-cm path length cuvette) via a mirror (M2) and a second beam splitter (B2), to the photodetector. The reference beam passes to the photodetector via mirror M1 and beam splitter B2. Mirror M1 is mounted on a piezoelectric transducer (PZT) for alignment purposes. The intensity of the signal at a point on the surface of the photodetector will be the result of the interference of the two recombined beams. Photodetectors, in

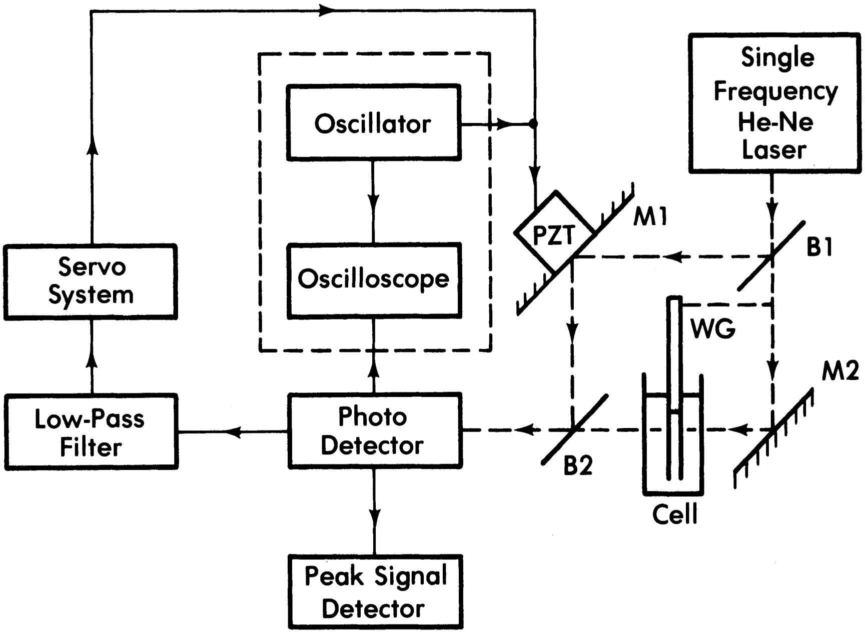


Figure 4 - Block diagram of PFLOH spectrometer.

general, are square-law responsive and thus the output signal will be proportional to the total incident intensity. Thus, any time-dependent phase or amplitude change in the signal beam will result in a variation in the intensity, which changes the output of the photodetector. For example, absorption of microwaves by a liquid sample will result in thermal expansion of the liquid and thus a change, or fluctuation, in its index of refraction. The fluctuation will cause an effective path length change for the signal beam. The phase change in the signal at the photo-detector is fed into a low-pass filter with a cut-off frequency of approximately 10 Hz. The resulting D.C. signal is amplified and used to drive the piezoelectric transducer which supports mirror M1. Thus, slow variations in the output signal result in compensating adjustment of the reference path length and alignment is maintained. Provided induced phase changes are small, the output signal will be linearly proportional to the magnitude of the disturbance and can be observed on the oscilloscope or the peak value recorded with a peak reading voltmeter.

Microwaves are introduced into the cuvette-contained sample by a specially constructed waveguide section depicted in Figure 5. This particular waveguide is dielectrically loaded to aid in matching microwave energy into the sample and to provide a uniform interface to the media. A 0.8 mm-wide slot running along the 5.72 X 2.85 mm

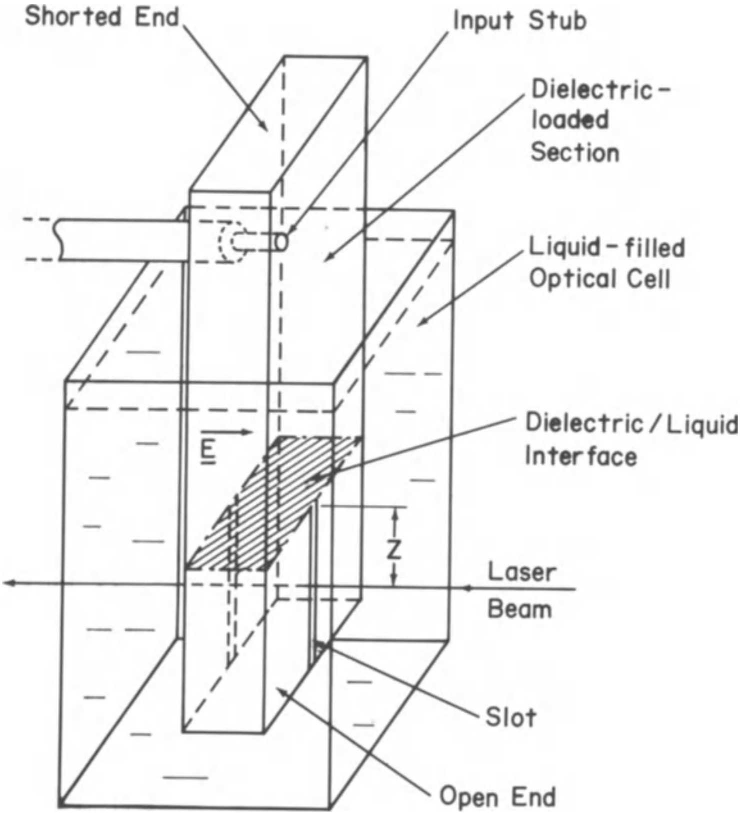

Shorted End
Input Stub
Dielectric-loaded Section
Liquid-filled Optical Cell
E
Dielectric / Liquid Interface
Z
Laser Beam
Slot
Open End

Figure 5 - Waveguide exposure system and sample holder. Waveguide dimensions = 0.57 X 0.28 cm.

waveguide allows passage of the laser beam across the narrow dimension of the waveguide parallel to the E field for the lowest order mode.

Measurements can now be made as a function of distance from the dielectric sample interface. The exponential decay ($\ell^{-\alpha x}$) of the signal with distance (x) determines the absorption coefficient (α) in the waveguide. For higher dielectric constant materials, the waveguide absorption coefficient is approximately the same as the absorption coefficients of the unbounded (infinite media) dielectric material. Movement in the laser beam was impractical. Therefore the waveguide section was moved with the aid of a stepping motor. The smallest increment of movement was 2 μm.

A two-millisecond pulse of RF was introduced into the waveguide and the resulting PFLOH signal was recorded as a function of distance

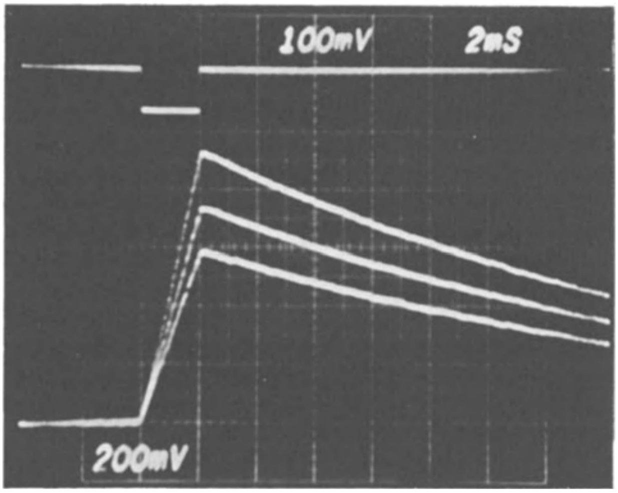


Figure 6 - Microwave induced PFLOH signals observed in de-ionized water. Timescale - 2 msec/div, upper trace - microwave pulse, lower traces - PFLOH signals taken at 0.3 millimeter increments down the waveguide.

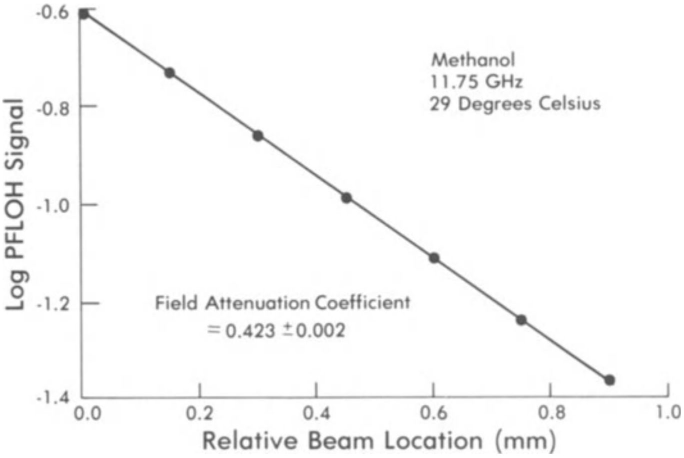


Figure 7 - Absorption of microwaves in methanol as a function of distance from the source in a waveguide.

down the guide. A scope trace of the detected RF signal and three resulting PFLOH signals detected at three positions in the guide is shown in Figure 6. A typical plot of PFLOH signal versus distance is presented in Figure 7 for observation in methanol at 11.75 GHz. Temperature of the sample was controlled to 0.1 degrees Celcius. Note, the calculated value of the attenuation coefficient obtained from the slope of the curve. The variability or repeatability of measurements was about ± 2%.

Data obtained for water in the 8 to 12 GHz frequency range are shown in Figure 8. Results agree with absorption coefficient values calculated from reported values of the complex dielectric constant. Calculations are made using the relationship.

$$\alpha = \left[1/2 \sqrt{\left(\left(\frac{\omega}{c}\right)^2 \varepsilon' - \left(\frac{n\pi}{a}\right)^2 - \left(\frac{m\pi}{b}\right)^2 \right)^2 + \left(\left(\frac{\omega}{c}\right)^2 \varepsilon' \tan\delta \right)^2} - \left(\left(\frac{\omega}{c}\right)^2 \varepsilon' - \left(\frac{n\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2 \right) \right]^{1/2}$$

and

$$\tan\delta = \frac{\sigma}{\omega\varepsilon_0\,\varepsilon'} + \frac{\varepsilon''}{\varepsilon'}.$$

Where ω is the angular frequency, C is the velocity of light, ε' and ε" are the real and imaginary parts of the dielectric constant, σ is the D.C. ionic conductivity, and a and b are the dimension of the waveguide. We have used the unusual convention of reporting the loss tangent as the sum of two terms emphasizing both conductive and dielectric losses.

Figure 9 shows a composite of data taken on a number of solutions. The heavy line through the lower portions of the data represents the average of the water absorption data presented in Figure 8. Note that the salts(saline, Ringer solutions and very impure DNA sodium salt)contributed to conductivity and increased absorption at the lower frequencies, but seemed to dilute the dielectric absorption properties of water at the higher frequencies.

The most interesting phenomenon is the considerable increase in absorption of the solution containing DNA. We extracted DNA from E. coli by conventional methods and resuspended the extraction at a concentration of what we then thought was 17 mg/ml but now believe to be much less. The D.C. conductivity of the DNA solution was measured and found to be a factor of 23 less than the D.C. conductivity of the saline solution. Thus, there is a 10 or 20% increase in the absorption because of a 1% by weight addition of DNA to the solution.

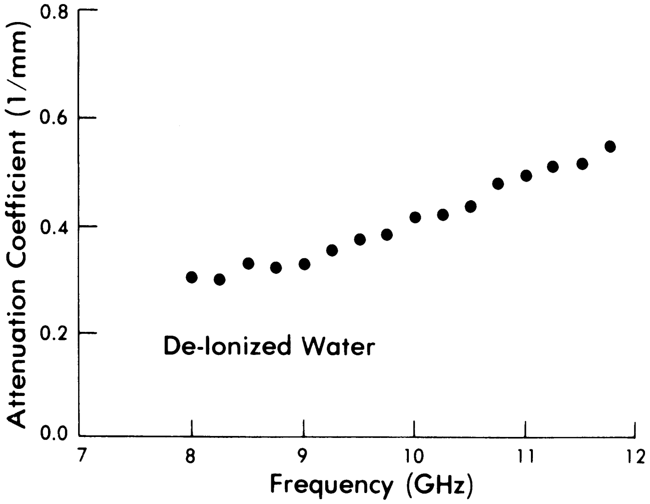


Figure 8 - Variation with frequency of the electric field attenuation coefficient of water in a waveguide of dimensions 0.572 X 0.286 cm at 29°C.

A theoretical group at Purdue University has predicted absorption modes in DNA in the radiofrequency range.[47,48] A recent letter in Biopolymers by Van Zant et al.[49] compared our results favorably with theoretical calculations suggesting the presence of critically damped longitudinal acoustic modes.

Critiques of our original findings have suggested that our results are due to either peculiarities of the newly developed PFLOH system or due to something like Maxwell-Wagner mixing theory. Our more recent results indicate that neither is the case and furthermore the results can not be interpreted on the basis of classical absorption theory.

We have since employed a second measurement technique, dielectrometry, and have improved our biochemical extraction methods which produced even more dramatic results. An automated network analyzer was available to us at the NCDRH laboratories. This device is schematically represented in Figure 10 and details of its operation have been described by Athey.[50] Basically, the device

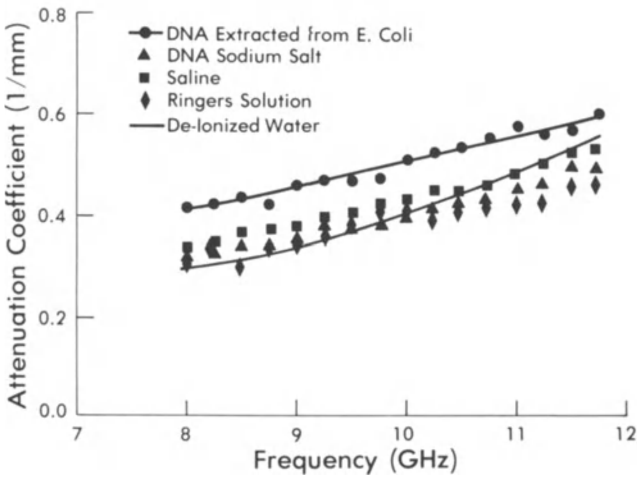


Figure 9 - Variation with frequency of the electric field attenuation coefficient of DNA, saline, Ringers solution and DNA sodium salt in a waveguide of dimensions 5.72 X 2.86 mm at 29°C.

measures the reflection coefficient using the indicated source and test sets under computer control. One method of determining the dielectric properties of a sample is the shorted line technique first described by Roberts and Von Hipple[51] in which the reflection coefficient is measured from a sample terminated in a short whose electrical length is 1/4 wavelength. The sample holder can be constructed from coaxial components in various lengths, as shown in Figure 11. The center conductor shown in Figure 11 determines the sample length. Similarly, sample holders can by constructed in a waveguide as illustrated in Figure 12.

In both cases, the output port of the analyzer is terminated in a Teflon-filled section. The mismatch due to the Teflon section is corrected by a calibration procedure. The Teflon provides the reference plane for the network analyzer as well as a flat interface for the liquid solution. The waveguide section shown in Figure 12 determines the sample thickness and is terminated by a short also shown.

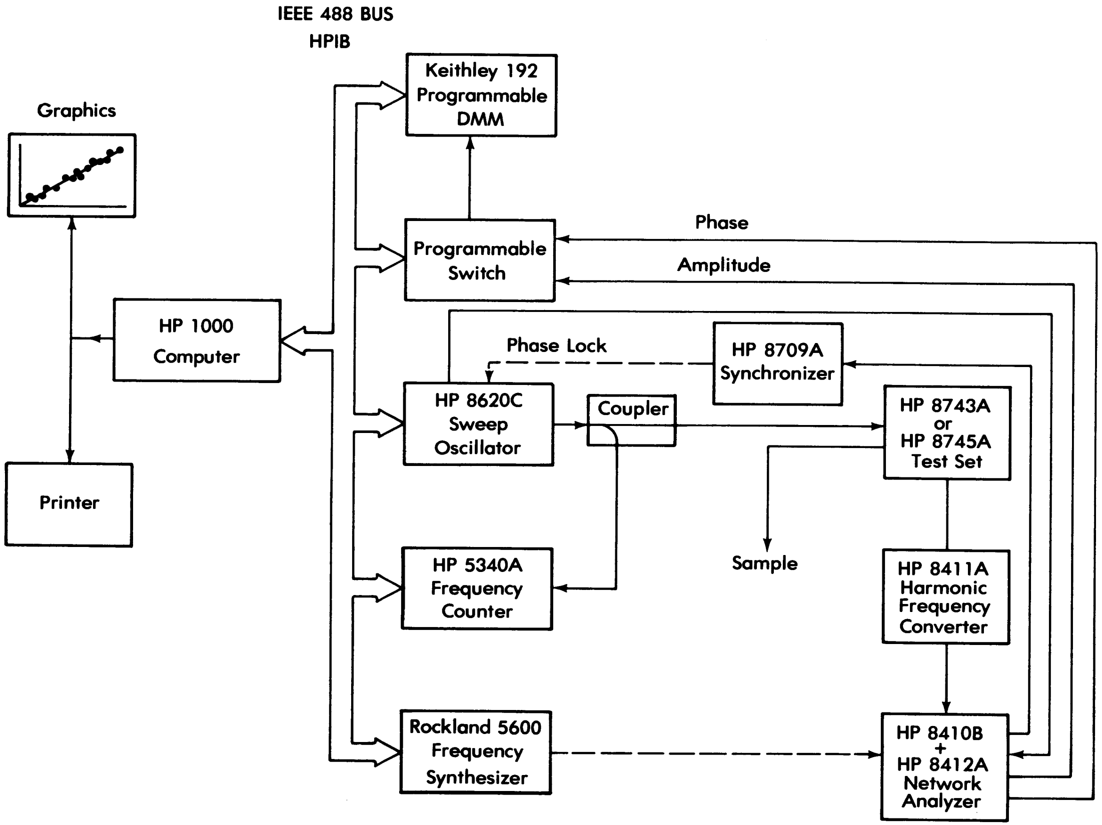

IEEE 488 BUS
HPIB
Graphics
Keithley 192 Programmable DMM
Programmable Switch
Phase
Amplitude
HP 1000 Computer
Phase Lock
HP 8709A Synchronizer
HP 8620C Sweep Oscillator
Coupler
HP 8743A or HP 8745A Test Set
Printer
Sample
HP 5340A Frequency Counter
HP 8411A Harmonic Frequency Converter
Rockland 5600 Frequency Synthesizer
HP 8410B + HP 8412A Network Analyzer

Figure 10 - Block diagram of computer controlled network analyzer. System is capable of measuring the complex reflection coefficient from the sample at specified frequencies and calculating the absorption coefficient.

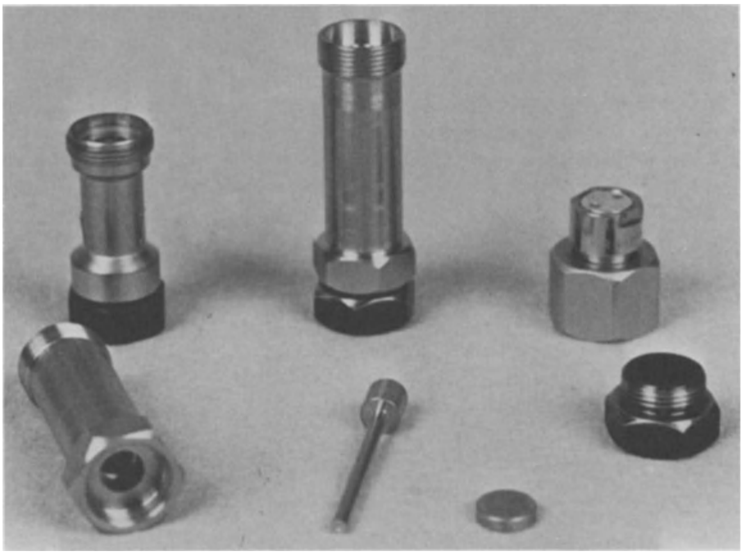


Figure 11 - Coaxial component sample holder for shorted line technique dielectric constant measurements.

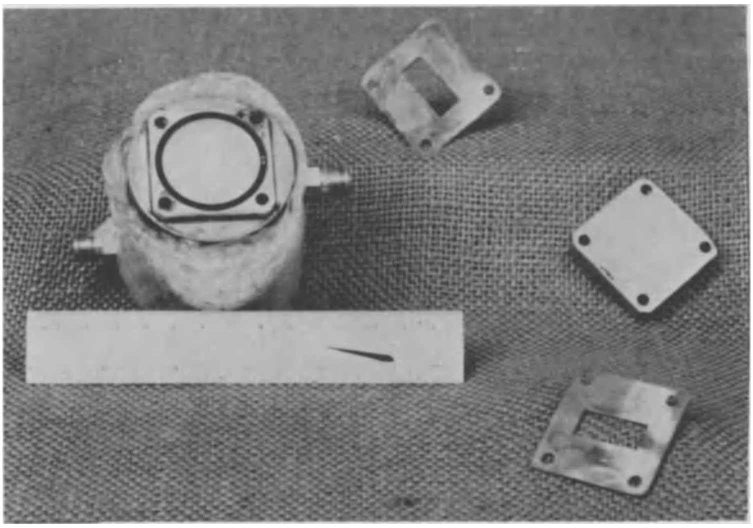


Figure 12 - Waveguide component sample holder used for shorted line technique dielectric constant measurements.

The complex dielectric properties of the sample can be determined from the reflection coefficient and one can then calculate the absorption coefficient using the equation previously shown. An indictation of the type of results (obtained with an estimated accuracy of 5%) is presented in Figure 13.

We attempted to improve the quality of the DNA used by purchasing purified E. coli DNA. Our initial results with the purchased material indicated no significant change in the absorption coefficient from the values obtained for the solvent. Being uncertain of the quality, we returned to our own extractions and again saw significant increases, confirming our earlier results. After standardizing the extraction procedure with one individual (Strickland) performing the extractions, we began to consistently (from preparation to preparation) observe the results presented in Figure 14.

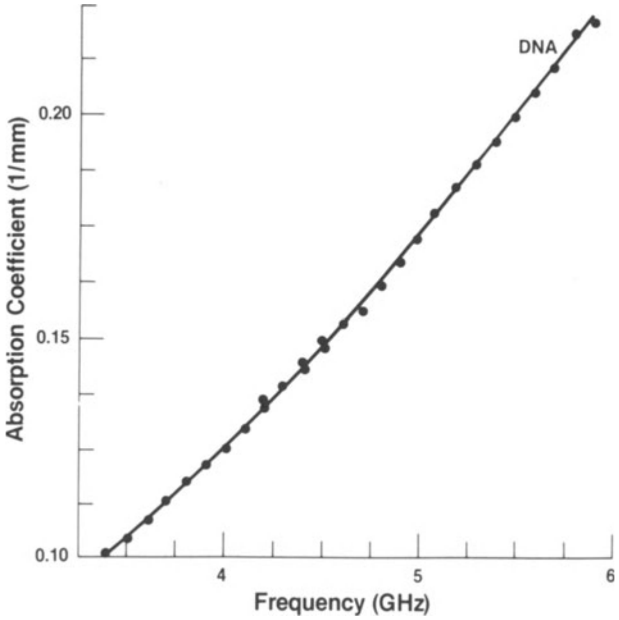


Figure 13 - Typical results showing the variation of the calculated absorption coefficient as a function of frequency using the coaxial sample holder.

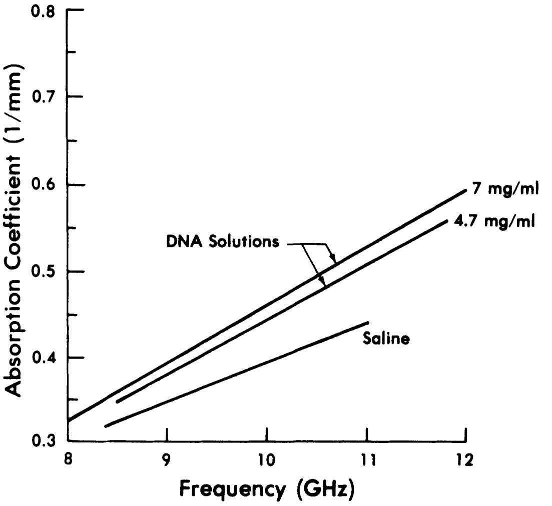


Figure 14 - Variation with frequency of the electric field attenuation coefficient for E. coli DNA in 9 mg/ml saline solution for 7 mg/ml and 4.7 mg/ml of DNA.

As in Figure 9, the frequency range is from 8 - 12 GHz. A very significant increase in the absorption of the DNA solution above the solvent is observed. The difference in solution and solvent increases with frequency as opposed to the slight decrease observed in the PFLOH results. This difference is believed to be principally, if not entirely, due to differences in sample preparations, although one might speculate as to differing responses of the two measurement methods.

The importance of sample preparation was brought dramatically home when we obtained the service of a visiting biochemist from Argentina, Jose Sagripanti. We asked Dr. Sagripanti to improve on our extraction technique and prepare pure samples of DNA in which we expected to see greater increases in the absorption coefficient. After the first measurement of his carefully prepared samples, we were surprised to observe very little increase in the absorption of the DNA solution over the solvent. Since the preparation method was more gentle, we assume the strands to be much longer. We had

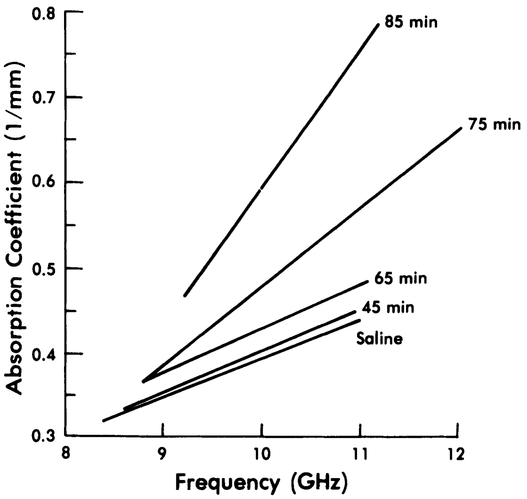


Figure 15 - Variation with frequency of the electric field attenuation coefficient of 1.75 mg/ml of E. coli DNA in 9 mg/ml saline solution for increasing time (indicated in figure) of exposure to 0.1 mg of DNAase and 10 mg/ml $MgCl_2$ per ml of solution.

incorrectly assumed that the longer strands would interact more strongly. Since this was not the case we presumed that treatment of the longer strand DNA with an enzyme, which broke the strands into smaller segments, would increase the absorption characteristics of the solution. The new DNA preparation was treated with 100 µg/ml of DNase I (an endonuclease) and the absorption properties were observed as a function of time. These results are depicted in Figure 15. As the enzyme acted on the DNA, presumably creating shorter and shorter chain lengths, the absorption coefficient increases to almost a factor of two above the solvent at eleven GHz. The DNA concentration is 1.75 mg/ml; again a very highly concentrated solution yet less than 0.2% DNA by weight. Not shown in this figure is the result of continued enzyme activity. As enzyme activity continues, the absorption will decrease back to the value of the solvent.

Resonant absorption dependent on chain length is certainly indicated. The results presented here are due to random lengths distributed about a mean which decreases with time.* We have continued our investigations using the "antenna probe" method described by Athey.[50] The preliminary results using this third measurement technique further substantiate our earlier findings.[45] To date we have not successfully isolated and measured the properties of a specific chain length. However, preliminary studies indicate that a chain length that is resonant in the X-band region contain less than 2000 base pairs.[52]

CONCLUSIONS

1. There is evidence to demonstrate that under certain exposure conditions RF energy may affect the genetic control of cell function, (gene expression and activation) including cell differentiation. Such effects should be investigated further because of far-reaching Biological implications and for potential beneficial applications in medicine and industry.

2. There is evidence to demonstrate that under certain exposure conditions RF energy may induce chromosomal aberrations both in somatic and germ cells. Further studies are necessary to establish dose-effect relationships and the dependence of these effects on frequency and modulation.

3. Both theoretical considerations and experimental data demonstrate that the DNA molecule reacts strongly with radiofrequency wave radiation. The induction of dipole moments along the chain producing longitudinal acoustic vibrational modes seems plausible. These interactions appear to be broad in frequency. The presence of hydrogen bonds along the DNA chain and histone bonds in the mammalian chromosomal system may provide the nonlinear interactive forces these could lead, as in the Davydov-Alpha-Helix-protein model, to the formation of robust solitary waves. The dispersive-vibrational modes reported here (on purified DNA and not the mamalian chromosomal system) are certainly energetic enough to drive a soliton generating process, provided that non-linear modes exist in the molecule and that the coupling to these modes is strong.

---

*It was suggested at the time of presentation by H.P. Schwann and H. Frohlich that the increased absorption may be in part due to bound water. This hypothesis should be experimentally tested.

REFERENCES

1. J. H. Heller and Teixeria-Pinto, A new physical method for creating chromosomal aberrations, Nature, 183:905 (1959).

2. J. H. Heller, Cellular effects of microwave radiation, in: "Biological Effects and Health Implications of Microwave Radiation," S. F. Cleary, ed., USDHEW, Rockville, MD, Report BRH/DBE 70-2, pp. 116-121 (1970).
3. A. A. Kapustin, M. T. Rudnev, G. I. Leonskaya, and G. I. Konobeyeva, Cytogenetic effects of a variable electromagnetic field in the superhigh frequency range (in Russian) Tsitologiya i Genetika, 5:400-402 (1976).
4. D. E. Janes, W. M. Leach, W. A. Mills, R. T. Moore, and M. L. Shore, Effects of 2450 MHz microwaves on protein synthesis and on chromosomes in chinese hamsters, Non-Ionizing Radiation, 1:125-134 (1969).
5. K. T. S. Yao and M. S. Jiles, Effects of 2450 MHz microwave radiation on cultivated rat kangaroo cells, in: "Biological Effects and Health Implications of Microwave Radiation," S. F. Cleary, ed., USDHEW, Rockville, MD, Report BRH/DBE 70-2: 123-133 (1970).
6. W. A. Stodolnik-Baranska, The effects of microwaves on human lymphocyte cultures, in: "Biological Effects and Health Hazards of Microwave Radiation," P. Czerski, K. Ostrowski, M. L. Shore, C. Silverman, M. J. Suess, and B. Waldeskog, eds., Polish Medical Publishers, Warsaw, 189-195 (1974).
7. K. T. S Yao, Cytogenetic consequences of microwave incubation of mammalian cells in culture, J Heredity, 73:133-138 (1982).
8. K. T. S. Yao, Microwave radiation-induced chromosomal aberrations in corneal epithelium of chinese hamsters, J Heredity, 69:409-412 (1978).
9. L. Cieciura, H. Hibner, L. Minecki, and G. Andry Szak, The influence of microwave radiation on the mitosis in root tips of the bean. Vicia Faba (in Polish) Med Pracy, 22:211-212 (1971).
10. K. M. Chen, A. Sammel, and R. Hoopingarner, Chromosomal aberrations of living cells induced by microwave radiation, Environmental Letters 6:37-46 (1974).
11. W. M. Leach, Genetic, growth, and Reproductive effects of microwave radiation, Bull. NY Academy Medicine, 56:249-257. (1980).
12. K. Hirschhorn, and M. M. Cohen, Induced chromosomal aberrations with special reference to man, in: "Comparative Mammalian Cytogenetics." K. Benirschke, ed., Springer, NY, 49 (1969).
13. W. C. Dewey, S. A. Sapareto, and D. A. Betten, Hyperthermia radio-sensitization of synchronous chinese hamster cells, relationship between lethality and chromosomal aberrations, Radiation Res., 76:48-59. (1978).
14. W. C. Dewey, A. Westron, H. H. Miller and H. Nagasawa, Heat induced lethality and chromosomal damage in synchronized chinese hamster cells treated with 5-bromodeoxy uridine, Int. J. Radiat. Biol. 20:505-520 (1978).
15. A. Westra, and W. C. Dewey, Variation of sensitivity to heat shock during cell cycle of chinese hamster cells in vitro. Int. J. Radiat. Biol. 19:467-477 (1971).

16. S. Mittler, Hyperthermia and radiation induced genetic aberrations in Drosophila melangaster. Mutat. Res. 59:123 (1979).

17. P. Smith, and H. O. Corwin, Temperature shock-induced specific visible mutations at the Dumpy locus of D. Melanogater. Mutat. Res. 13:345-351 (1971).

18. E. Manikowska, J. M. Luciani, B. Servantie, P. Czerski, J. Obrenovitch and A. Stahl, Effects of 9.4 GHz exposure on meiosis in mice.

14. W. C. Dewey, A. Westron, H. H. Miller and H. Nagasawa, Heat induced lethality and chromosomal damage in synchronized chinese hamster cells treated with 5-bromodeoxy uridine, Int. J. Radiat. Biol. 20:505-520 (1978).

15. A. Westra, and W. C. Dewey, Variation of sensitivity to heat shock during cell cycle of chinese hamster cells in vitro. Int. J. Radiat. Biol. 19:467-477 (1971).

16. S. Mittler, Hyperthermia and radiation induced genetic aberrations in Drosophila melangaster. Mutat. Res. 59:123 (1979).

17. P. Smith, and H. O. Corwin, Temperature shock-induced specific visible mutations at the Dumpy locus of D. Melanogater. Mutat. Res. 13:345-351 (1971).

18. E. Manikowska, J. M. Luciani, B. Servantie, P. Czerski, J. Obrenovitch and A. Stahl, Effects of 9.4 GHz exposure on meiosis in mice. Experientia 35:388-390 (1979).

19. E. Manikowska-Czerska, P. Czerski, and W. M. Leach, Changes in metaphase counts, meiotic translocations and univalency following 2.45 GHz exposure of male mice. 4th Annual Session BEMS, abstracts, Los Angeles, p. 85. (1982).

20. E. Manikowska-Czerska, P. Czerski, W. M. Leach, Microwave radiation exposure induces translocations and univalents during spermatogenesis in mice. J Cell Biol 95:301 (1982).

21. E. Manikowska-Czerska, P. Czerski, and W. M. Leach, Effects of 0.915 and 9.4 GHz CW Microwaves on meiosis in male mice. 5th Annual Scientific Session BEMS, Abstracts, Boulder, p. 57 (1983).

22. E. Manikowska-Czerska, P. Czerski, and W. M. Leach, Dominant lethal testing after exposure of mice to 0.915 GHz microwaves. 5th Annual Scientific Session BEMS. Abstracts, Boulder, p. 57 (1983).

23. W. M. Varma and E. A. Traboulay, Jr. Evaluation of dominant lethal test and DNA studies in measuring mutagenicity caused by non-ionizing radiation, in: "Biological Effects of Electromagnetic Waves," C. C. Johnson, and M. L. Shore, eds., Rockville, MD, Food and Drug Administration (HEW Publication, FDA, 77-8010), Vol. 1:386-396 (1976).

24. W. M. Varma, E. L. Dage, and S. R. Joshi, Mutagenicity induced by nonionizing radiation in swiss male mice, in: "Biological Effects of Electromagnetic Waves," C. C. Johnson and M. L. Shore, eds., Rockville, MD, Food and Drug Administration (HEW Publication, FDA, 77-8010), Vol. 1:397-405 (1976).

25. S. N. Goud, M. V. Usha Rani, P. P. Reddy, M. S. Rao, and V. K. Saxena, Genetic effects of microwave radiation in mice. Mutat. Res. 103:39-42 (1976).

26. E. Berman, H. B. Carter, and D. House, Tests of mutagenesis and reproduction in male rats exposed to 2450 MHz (CW) microwaves. Bioelectromagnetics 1:65-76 (1980).

27. E. F. Oakberg, A new concept of spermatogonial stem-cell renewal in the mouse and its relationship to genetic effects, Mutation Res., 11:1-7 (1971).

28. J. L. Oud, J. H. de Jong, D. G. de Rooij, Sequential analysis of meiosis in the male mouse using a restricted spermatocyte population obtained by a hydroxyurea/triaziquone treatment, Chromosoma (Berlin), 71:237-248 (1979).

29. E. Aarholt, E. A. Flinn, and C. W. Smith, Magnetic fields affect the lacoperon system. Physics in Med. and Biol. 27:606-610 (1982).

30. W. M. Generoso, K. T. Cain, S. W. Huff, and D. G. Gosslee, Heritable translocation test in mice, in: "Chemical Mutagens, Principles and Methods for Their Detection." A. Hollaender and F.J. de Serres, eds., Plenum Press New York and London. Vol. V, pp. 55-78 (1978).

31. A. Z. Smolanskaya and R. L. Vilenskaya, Effects on millimeter band electromagnetic radiation on the functional activity of certain genetic elements of bacterial cells, in: "Scientific Sessions of the Division of General Physics and Astronomy," USSR Academy of Sciences January 17-18, 1973.

32. Webb, J. Sydney, Factors affecting the induction of lambda prophages by millimeter microwaves. Physics Letters 73A:145-148 (1979).

33. R. Goodman, C. A. L. Bassett, and A. S. Henderson, Pulsing Electromagnetic fields induce cellular transcription. Science 222:1283-1285 (1983).

34. F. Kremer, C. H. Koschnitzke, L. Sants, P. Quick, and A. Poglitsch, The nonthermal influence of millimeter wave radiation on the puffing of giant chromosomes. 5th Annual Scientific Session of BEMS. Abstracts, Boulder, p. 56 (1983).

35. A. Chiabrera, M. Hinsenkamp, A. A. Pilla, J. Ryaby, D. Ponta, A. Belmont, F. Beltrame, M. Grattarola, and G. Nicolini, Cytofluorometry of electromagnetically controlled cell dedifferentiation. J Histo chem Cytochem 27:375-381 (1979).

36. D. J. McRee, G. MacNichols, and G. K. Livingston, Incidence of sister chromatid exchange in bone marrow cells of the mouse following microwave exposure. Radiation Res. 85:340-348 (1981).

37. S. Baranski, H. Debiec, K. Kwarecki, and T. Mezykowski, Influence of microwaves on genetical processes of Aspergillus nidulans. J of Microwave Power 11:146-151 (1976).

38. C. F. Blackman, S. G. Benane, C. M. Weil, and Y. S. Ali, Effects of nonionizing electromagnetic radiation on single-cell biologic systems. Annals NY Acad Sci 247:352-366 (1975).

39. M. Dardalhon, D. Averbeck, and A. J. Bertand, Determination of a thermal equivalent of millimeter microwaves in living cells. J Microwave Power 14:307-312 (1979).
40. S. K. Dutta, W. H. Nelson, C. F. Blackman, and D. J. Brusick, Lack of microbial genetic responses to 2.45 GHz CW and 8.5 to 9.6 GHz pulsed microwaves. J Microwave Power 14:275-280 (1979).
41. S. Mitler, Figure of 2 and 10 meter radiowaves to induce genetic damage in Drosophila melanogasten. Environmental Res. 11:326-330 (1976).
42. T. L. Pay, E. G. Beyer, and C. F. Reichelderfer, Microwave effects on reproductive capacity and genetic transmission in Drosophila melanogaster. J Microwave Power 7:75-82 (1972).
43. M. L. Swicord, G. S. Edwards, S. L. Sagripanti, and C. C. Davis, Chain length dependent microwave absorption of DNA. Submitted for publication (1983).
44. C. C. Davis and M. L. Swicord, Studies of microwave absorption in liquids by optical heterodyne detection of thermally induced refractive index fluctuations. Radio Science, Vol. 17, #5S, September-October, pp. 85S-94S (1982).
45. M. L. Swicord, C. C. Davis, Microwave absorption of DNA between 8 and 12 GHz. Bioelectromagnetics 4:21-42 (1982).
46. M. L. Swicord and C. C. Davis, An optical method for investigating the microwave characteristics of DNA and other biomolecules in solution. Bioelectromagnetics 4:21-42 (1983).
47. E. W. Prohofsky, Effects of microwave absorption by the DNA double helix, in: "The Mechanisms of Microwave Biological Effects," A report on a workshop at the University of MD, L. S. Taylor and A. Y. Cheung, eds., May 14-16, p. 7 (1979).
48. M. Kohli, W. N. Mei, E. W. Prohofsky, Van Zandt, Calculated microwave absorption of double-helical B-conformation Poly (dG) Poly (dC). Biopolymers 20:853-864 (1981).
49. L. L. Van Zandt, M. Kohli, E. W. Prohofsky, Absorption of microwave radiation by the double helix in aquo. Biopolymers 21: 1465-1468 (1982).
50. T. W. Athey, M. A. Stuchly, S. S. Stuchly, Measurement of radiofrequency permittivity of biological tissues with an open ended coaxial line: Part 1. IEEE Trans., MTT-30:82-92 (1982).
51. A. Roberts, Von Hipple, Jr., A new method for measuring dielectric constant and loss in the range of centimeter waves. J of Applied Physics 17:610-616 (1946).
52. G. S. Edwards, M. L. Swicord, C. C. Davis, Microwave absorption characteristics of highly purified E. coli DNA. 5th Annual Scientific Session BEMS, Boulder, p. 3 (1983).

NONTHERMAL RESONANT ACTION OF MILLIMETER MICROWAVES ON YEAST GROWTH

F. Keilmann and W. Grundler

Max-Planck-Institute fur Festkorperforshung
7000 Stuttgart 80, F.R. Germany

Gesellschaft fur Strahlen- und Umweltforschung
8042 Neuherberg, F.R. Germany

ABSTRACT

This talk describes experimental results which prove that microwaves can interact with biological cells in a nonthermal manner. It is shown that observed microwave effects differ strongly from those obtained by applying other Types of energy instead of microwaves at equivalent rates. The two most intriguing aspects of the interaction are that there exists an intensity threshold above which the effects appear, and that the effects are different for minute changes of microwave frequency. A step-like intensity dependence is paired with a resonant frequency dependence. The latter is described by a quality factor of about 10^4. The effects are as yet unexplained. The locus of interaction is not known. It is however remarkable that both an intensity threshold and a resonant behaviour are predicted in H. Frohlich's hypotheses of coherent excitations in biology.

INTRODUCTION

The possibility of microwave biological actions beyond those based simply on microwave heating has been much debated. This question is of fundamental interest (of course there is also the aspect of radiation protection (1). We shall address, as in previous papers (2-4), experimental techniques and results obtained with microwaves in the millimeter range, i.e. frequencies above 30 GHz.

Historically, nonthermal millimeterwave bioeffects were independently discovered by Webb (5) and Devyatkov (6) and their collaborators. Validity of these results has been questioned since then, for various reasons. For example, a criticism was that the reports did not contian sufficient detail to rule out standing wave resonances of the transmission lines. Attempts to reproduce Webb's results succeeded in one case (7) but failed in another (8).

Further work (see references in 11) was done in the USSR, USA and FRG. Studies in USA did not answer positively the question of nonthermal millimeterwave bioeffects. On the other hand, a considerable number of positive results were compiled in the USSR Academy of Sciences but it was only recently that we became aware of these studies (9). In Germany (10-13), activity was spurred by the theoretical work of Frohlich (14), to be discussed in detail at this meeting.

Since our original discovery of unexpectedly sharp resonances in the microwave spectrum for responses in yeast cell growth (10), we have concentrated on corroborating these findings, with positive results (11). We are grateful to L. Genzel and the Deutsche Forschungsgemeinschaft for their support. Regarding future activities, I take this opportunity to announce an international symposium devoted solely to experimental procedures in studies of low-level millimeterwave bioeffects, to be held in Herrsching, September 4-6, 1983.

MEASUREMENT OF YEAST GROWTH

Yeast cells which are a few micrometers in size were suspended in growth medium. The suspension was constantly stirred to avoid settling. With an optimum nutritional environment, the cells coupled roughly once every hour. We observed the growth by recording optical density or turbidity (15) in a spectrometer at 550 nm. The output signal V_{OD} was amplified logarithmically and continuously recorded versus time. A straight line indicates exponential growth (yeast cells are not synchronized). The slope of this line was read with ± 1% accuracy and gives directly the exponential growth rate $= t \ln V_{OD}$. Two photometers were employed for simultaneous irradiation and control experiments. Without irradiation, the fluctuations between both measurements stayed within ± 4%. This allowed measurement with great significance of microwave induced alterations in growth rate of roughly 10%.

MICROWAVE-INDUCED TEMPERATURE RISE

The temperature rise induced at an irradiation intensity of 10 mW/cm can be as large as 5.5%C. This was estimated (3,4) for the cases where only two cooling mechanisms contribute, namely infrared emission and passive air convection. The experiments (5-13) used incident intensities between 10 W/cm and 10 mW/cm . This required additional cooling techniques at the high end of this intensity range, to ensure that microwave-induced temperature increments stayed below some insignificant value, e.g. 0.1 or 0.01 C. In practice, this goal was reliably achieved by heat conduction in water. All that was necessary was to immerse the specimens in water, which was also heat-sinked, preferably by a good heat conductor such as sapphire (3). In our yeast experiment this heat-sinking was designed for a measurable but small temperature rise of the order of 0.3 C, since this allowed a quite direct measurement of the absorbed microwave energy.

IMPOSSIBILITY OF MICROSCOPIC HOT SPOTS

It has been argued that this microwave energy absorbed locally in some cellular of subcellular unit might lead to significant local overheating. This can be rigorously ruled out (3,4) by two findings. First, there exist upper limits on the specific absorption rate or power absorption per unit volume, as determined from spectroscopic considerations. Secondly, conduction cooling increases in efficiency as the exposed volume shrinks. An exception could only occur in the hypothetical case that the absorbing unit was somehow thermally isolated from its surroundings.

TRANSMISSION LINE RESONANCES

The high coherence of common microwave sources leads to standing-wave problems typical of most microwave applications: even small reflections at source and target suffice for build-up of transmission line resonances. One consequence of importance to our work is that even small relative changes in either microwave frequency or in length of transmission line drastically alter power deposited in the target (3,4). If paired with a thermal microwave effect, for example, this can lead to artefactual resonances, with roughly 100 MHz spacing with usual transmission line lengths (3,4). In our yeast experiment these difficulties were dealt with by (i) using calorimetric determination of absorbed microwave power as described above, and (ii) by using bidirectional couplers to measure incident and reflected power.

EXPERIMENTAL RESULTS

A long series of repeated yeast growth experiments (11) resulted in the response spectrum shown in Fig. 1. In each experiment the frequency was stabilized to an absolute accuracy within ± 1 MHz. Statistical calculations (11) showed an over-

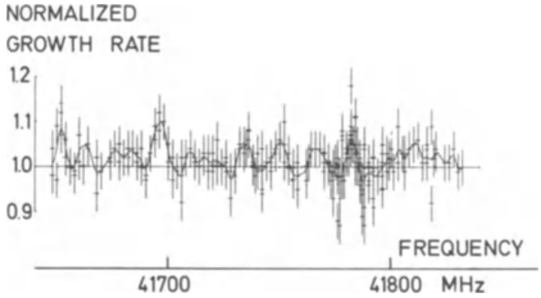


Figure 1. Nonthermal effect of microwave irradiation on the growth rate of yeast in suspension, versus the microwave frequency. The curve is obtained by a single three point smoothing procedure.

whelming probability that this spectrum did not originate from random fluctuations, but revealed a truly frequency-depdendent microwave effect. The possibility of artefactual resonances (originating in standing wave resonances of the waveguide system) was not only ruled out by theoretical considerations noted above. In addition, two quite different microwave antennae were employed (11) to obtain the data shown in Fig. 1. If standing wave resonances had exerted an influence, the frequencies of the growth rate resonances would have shifted. This did not occur.

Experimental points in Fig. 1 were not obtained at a constant power setting. The absorbed power was constant in a given experiment but varied between runs over a range from 6 to 38 mW. In an attempt to determine power dependence of the microwave effect, we replotted (11) all experimental points vs. power, irrespective of frequency. This plot did not reveal any correlation, an observation which by itself speaks for the nonthermal nature of the effect. On the other hand we noted a very interesting flat power dependence when the frequency fixed was kept constant (at 41782 MHz): the effect remained constant irrespective of absorbed power over the range 4 to 22 mW. This denotes the existence of an intensity threshold. However, its value cannot be accurately determined because of the inhomogeneous intensity distribution in the yeast suspension.

OUTLOOK

The narrow width of resonance observed in our experiment, together with a flat power dependence prove the existence of a nonthermal microwave sensitivity inherent in yeast cells. This enhances evidence in favor of similar reports from other biological specimens (5-13). The question of generality of such a sensitivity in biology can be posed.

The significance of such a sensitivity can only be answered when more is known of possible interaction mechanisms. Far-reaching consequences for our understanding of biology would follow in the event that future experiments support the basic conjecture of Frohlich (14) in further detail. Microwave resonance would then indicate an active role for long-range interactions, complementing in living cells previously known short-range chemical forces and diffusional motions.

We are now trying to answer experimentally a series of physical and biological questions. These encompass shifting the resonances by application of external parameters and fields or by variation of the biological species. Preliminary tests with yeast cells fixed on agar (16) seem to show no shift in resonant frequencies. This method allows both a well-defined homogeneous irradiation intensity. By microscopic inspection of individual cells, it will be possible to study dependence of these microwave effects on the cell cycle phase.

REFERENCES

1. With regard to microwave safety driving public anxieties (P. Brodeur, The Zapping of America, Norton, New York, 1977) from a scientiffically weak basis, this probably hinders the advancement of science. Thus, such an emotional atmosphere generates preconcpetions and may even lead to attempts to "prove" the nonexistence of any nonthermal biological microwave effects.
2. F. Keilmann, in:"Coherence in Spectroscopy and Modern Physics," F.T. Arecchi, R. Bonifacio and M.O. Scully, (eds.) Plenum, New York, pp. 347-360, (1978).
3. F. Keilmann, Collective Phen. 3:169-180 (1980).
4. F. Keilmann, in: "Biological Effects and Dosimetry of Nonionizing Radiation, M. Grandolfo, S.M. Michaelson and A. Rindi, (eds.), Plenum, New York, pp. 283-297, (1983).
5. S.J. Webb and D.D. Dodds, Nature 218:374, (1968); S.J. Webb and A.D. Both, Nature 222:1199, (1969).
6. N.D. Devyatkov et al., Sov. Phys. Usp. 16:568, (1974).
7. A.J. Berteaud, M. Dardalhon, N. Rebeyrotte and D. Averbeck, C.R. Acad. Sc. Paris 281:D843, (1975).
8. C.F. Blackmann, S.G. Benane, C.M. Weil and J.S. Ali, Ann. New York Acad. Sci. 247:352, (1975).
9. N.D. Devyatkov, (ed.) "Nonthermal effects of Millimeter wave Irradiation," Academ. Sci. USSR, Inst. Radiotech. Electrotech., Moscow, in Russian (1981).
10. W. Grundler, F. Keilmann and H. Frohlich, Phys. Lett. 62A:463; W. Grundler and F. Keilmann, Z. Naturforsch. 33c:15 (1978).
11. W. Grundler, F. Keilmann, V. Putterlik, L. Santo, D. Strube and I. Zimmermann, in: "Coherent Excitations in Biological Systems," H. Frohlich and F. Kremer, (eds.), Springer Verlag, Heidelberg, (1983).
12. F. Kremer, C. Koschnitzke, L. Santo, P. Quick and A. Poglitsch, ibid., pp. 13-24.
13. G. Nimtz, ibid., pp. 47-55.
14. H. Frohlich, Int. J. Quant. Chem. 2:641, (1968); H. Frohlich, Nature, 228:1093 (1970); H. Frohlich, in: "Advances in Electronics and Electron Physics," L. Marton and C. Marton, (eds.), Academic Press, New York, Vol. 53, pp. 85-152, (1980).
15. F. Keilmann, D. Bohme and L. Santo, Appl. Envir. Microbiology 40:458 (1980).
16. W. Grundler, unpublished.

EFFECTS ON NERVOUS TISSUE OF EXPOSURE TO ELECTROMAGNETIC FIELDS

Hans-Arne Hansson

Institute of Neurobiology
University of Gothenburg
Post Office Box 33031
S-400 33 Gothenberg, SWEDEN

ABSTRACT

Results obtained *in vitro* or by biochemical, pharmacological and behavioral studies suggest that nervous tissue is affected by exposure to extremely low frequency electromagnetic fields. The present report summarizes research performed during the last few years mainly using light and electronmicroscopic and immunohistochemical methods.

Long term exposure of animals, i.e. rabbits, pigs, rats and mice, at moderate to high intensity electromagnetic fields (50 or 60 Hz) results in significant reactive changes in the nervous tissue. The animals were exposed either during their first four weeks after birth or during their fetal life and then up to about 4 weeks of age.

We chose the cerebellum as the structure to be studied in detail because it is very regularly organized and shows only limited variation in structure between, adjacent cells of the same type.

Both the cell body and the radiating processes of Bergmann's glial cells showed extensive changes as did astrocytes in the granular layer and white matter. Lamellar bodies, i.e. rebuilt endoplasmic reticulum, could be seen in large number in Purkinje nerve cells from rabbit, but not in specimens from rodents. We observed

cytoskeletal alterations in most Purkinje nerve cells. Thus, in principle, similar changes were induced in all animal species examined.

Immunohistochemical methods for demonstration of glial reactions enable us to confirm that central nervous tissue is significantly affected by long term exposure to power line frequency electromagnetic fields. The mechanisms causing the described changes are not known but are certainly non-thermal.

INTRODUCTION

During last years increasing evidence has been presented indicating significant biological effects on animals and humans exposed to extremely low frequency (ELF) electromagnetic fields (Sheppard and Eisenbud, 1977; Adey, 1981; Becker and Marino, 1982; Sheppard, 1983).

The nervous system may be affected both by acute and long term exposure not only in vitro but also in vivo. The circadian rhythm and its related metabolic processes are further significantly altered due to chronic exposure to low levels of power frequency electric fields (Wilson et al., 1982).

We have previously demonstrated a strongly increased frequency of nerve cell changes in the cerebellum and other areas in the brain of rabbits exposed to 50 Hz electromagnetic fields (Hansson, 1981 a, b). However, the significance from structural and functional points of view of the various neuronal effects described were not known.

In the mammalian nervous system there are intimate interactions between nerve cells and their surrounding and supporting glial cells. Any significant affection of either cell type is likely to cause changes in the other interacting cells

Therefore, we decided to investigate in detail with immunochemical, immunohistochemical, light and electron microscopic techniques whether the most frequent supporting cell type in the central nervous system, i.e. astrocytes, showed reactive changes. Specimens from several different species were examined in order to elucidate if a common reaction pattern could be demonstrated.

A prerequisite for being able to resolve even discrete glial changes in nervous tissue is to examine regions which are organized in an exact way, allowing easy identification of various cell types. Furthermore, only limited variability should occur between adjacent cells of the same type and in the same position. The cerebellum offers a unique system fulfilling these basic criteria (cf. Palay and Chan-Palay, 1974).

The observations to be reported as results of cooperative studies with several different laboratories support our initial working hypothesis that exposure of animals to power line electromagnetic fields of moderate intensities for longer time periods significantly affects the nervous system as reflected by reactive glial changes.

RABBITS

Lop-eared and albino New Zealand rabbits were reared either outdoors in a 400 kV substation or indoors in a laboratory. They were kept in plastic cages which were checked not to interfere with the electric fields and had free excess to water and food. We have examined brains from animals exposed either during both their fetal development and first 4-6 weeks postnatally or from birth up to the age of 6 weeks. The dam and her offspring, reduced to 6 or 7 at birth, were kept in the same cage. All experiments were done at 14 kV/m, 50 Hz and the exposure was only interrupted for one hour per day to enable animal care. In parallel, two different control experiments were performed, i.e. rabbits were kept in cages surrounded by grounded metallic nets or beyond measurable electric fields. At least twelve brains from animals (i.e. two groups) from each exposure system and their corresponding controls were examined.

The rabbits exposed to electric field outdoors did not gain weight as their corresponding controls (Hansson, 1981 a, b). However, those exposed to the same electric field intensity, i.e. 14 kV/m, indoors did not differ significantly from the controls with regard to growth rate or weight. We therefore presume that the extra stress on the exposed animals due to the fairly rough climate added by the outdoor exposure is a main cause for the weight difference.

Rabbit brains for light and electron microscopic examination were fixed by transcardial perfusion with buffered glutaraldehyde, postfixed in osmium tetroxide and then embedded in epon.

Previously reported results (Hansson, 1981 a, b) could be confirmed in these later series of experiments. Briefly, the Purkinje nerve cells showed reduced diameters as compared to those in the controls. A striking observation was the large number of lamellar bodies, i.e. rebuilt endoplasmic reticulum, present in the exposed but rare or absent in the controls. This difference is consistent even if the endoplasmic reticulum and cytoskeleton of rabbit Purkinje nerve cells are liable to rapid changes if not adequately fixed. Our results indicate large numbers of lamellar bodies only are observed in Purkinje nerve cells with extensive alterations of their cytoskeleton, presumably reflecting structural and, probably, functional interconnection. Furthermore, glial reactions were obvious when comparing the structure and relative surface area of astrocytes in specimens from exposed rabbits and those from controls.

Immunohistochemical studies on the distribution of the S-100 protein, a marker for protoplasmic astrocytes, revealed altered distribution of the Bergmann glial cells, interposed between adjacent Purkinje nerve cells. The cell bodies of these astrocytes further were more spherical and showed only few processes emerging from the perikaryon towards the leptomeningel covering of the cerebellum of exposed rabbits as compared to controls. The S-100 immunoreactive material was more concentrated in the astrocytic nuclei and perinuclear region in sections from exposed rabbits as compared to the strictly cytoplasmic, diffuse distribution of this marker in the controls.

The gliofibrillary acidic protein (GFA) was used as a marker for fibrous astrocytes as well as for the superficial parts of the glial radiating fibres in cerebellar section, using immunohistochemical techniques. The exposed rabbits were observed to have an increased number of astrocytes both in the granular cell layer and in the white matter of cerebellum as compared to the controls.

It may therefore be concluded that long term exposure of growing rabbits to power line electromagnetic fields of moderate intensity results in both neuronal and neuroglial reactive changes. These results are further supported by the preliminary results from immunochemical studies on the concentration of S-100 related both to cerebellar protein amount and wet weight, indicating increased amount of the glial protein in the lateral parts of cerebellum from exposed animals (unpublished observations).

RATS

In order to further elucidate whether the neuronal and neuroglial changes observed in the cerebellum and other parts of the brain of rabbits exposed for long time periods to moderate to high ELF electric fields could be reproduced in other animal species, a study was performed using albino rats. The dams and their offsprings were exposed for four weeks to powerline type (60 Hz) electric fields (14 kV/m) starting from birth of the sib Thereafter, the small rats, exposed and controls, were examined by light and electron microscopic techniques as well as fluorescence microscopy using the glial proteins S-100 and GFA as markers.

The distribution of GFA immunoreactive astrocytes turned out to be a useful indicator of effects on the nervous system induced by long term exposure to electric fields. The scattered fibrous astrocytes observed in the granular layer of the cerebellum in the controls, sham-exposed and cage controls (Fig. 1a) became markedly increased in both number and size (Fig. 1b). The most prominent increase was noted in the lateral hemispheres of cerebellum (Figs. 1a, b). In the posterior vermis the increase in number of immunoreactive cell bodies increased with a factor of at least two (Fig. 2). Furthermore, the dimensions of their perikarya increased as did the number of processes encircling the many small nerve cells in the granular layer. The astrocytes in the white matter also showed similar changes, i.e. increase with regard to both their number and size (Fig. 2b). Studies are in progress in order to evaluate whether similar changes also take place in other regions of the nervous system.

In parallel, other sections were examined by immunofluorescence microscopic methods for the distribution of S-100 immunoreactive cells. The Bergmann glial cells (Fig. 3) became beautifully labelled. It was confirmed that there was an increased number of such cells in specimens from exposed rats as compared to controls. Similar to the changes observed in the cerebellum of exposed rabbits, the S-100 reactivity was concentrated to the nuclei and immediate perinuclear cell region, reflecting that the number of radiating processes from each cell body was decreased (Fig. 3b).

Light and electron microscopic examination of epon-

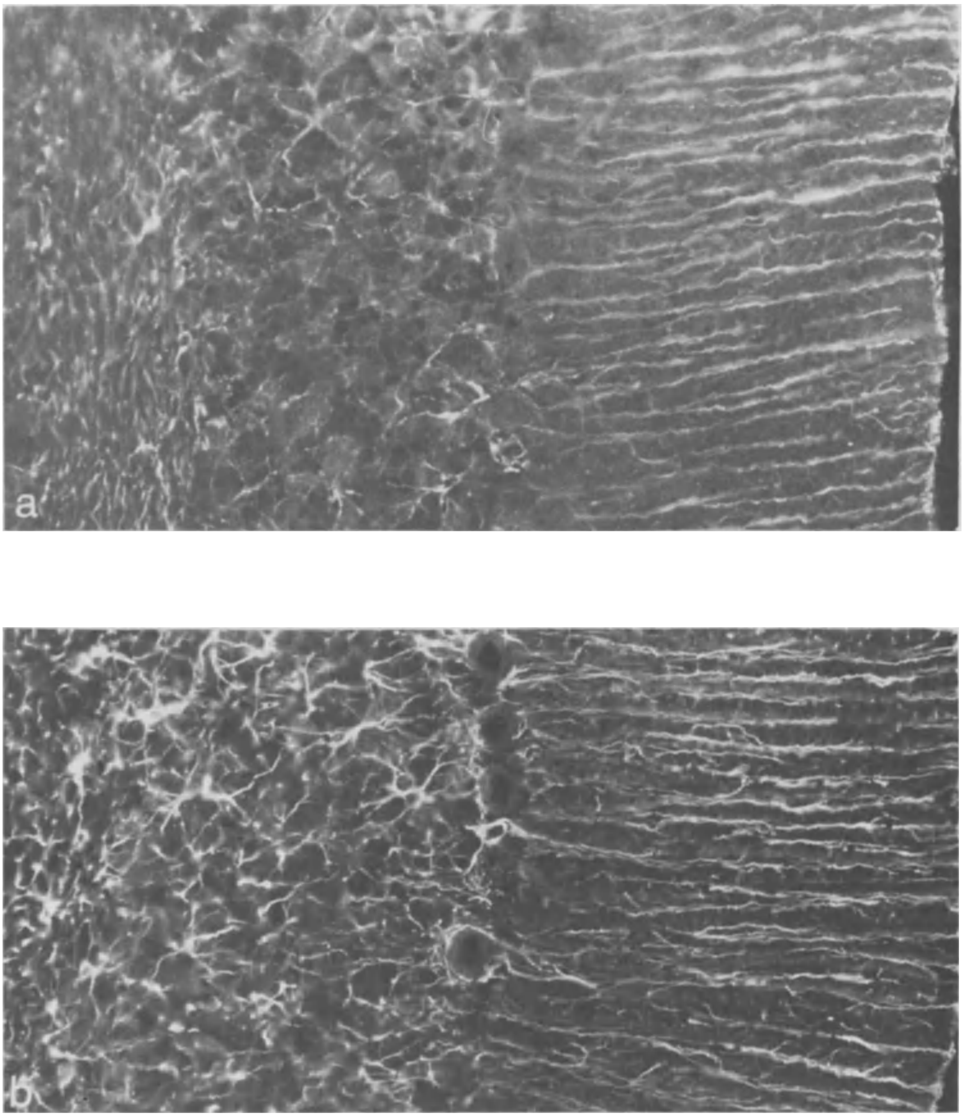


Fig. 1: Fluorescence micrograph of cerebellum from rat, processed for demonstration of gliofibrillary acidic protiens(GFA). 1a; Cage control. Scattered astrocytes with delicate processes seen adjacent to Purkinje nerve cells (P). Bergmann's radial glial(B) processes extend through the molecular layer. 1b; Exposed for 4 wks. 14 kV/m 60 Hz. Increased number of reactive astrocytes (arrow), partly encircling Purkinje nerve cells(P). Note the extensive meshwork of astrocytic processes in the granular cell layer (arrowhead). Lateral hemispheres of cerebellum.

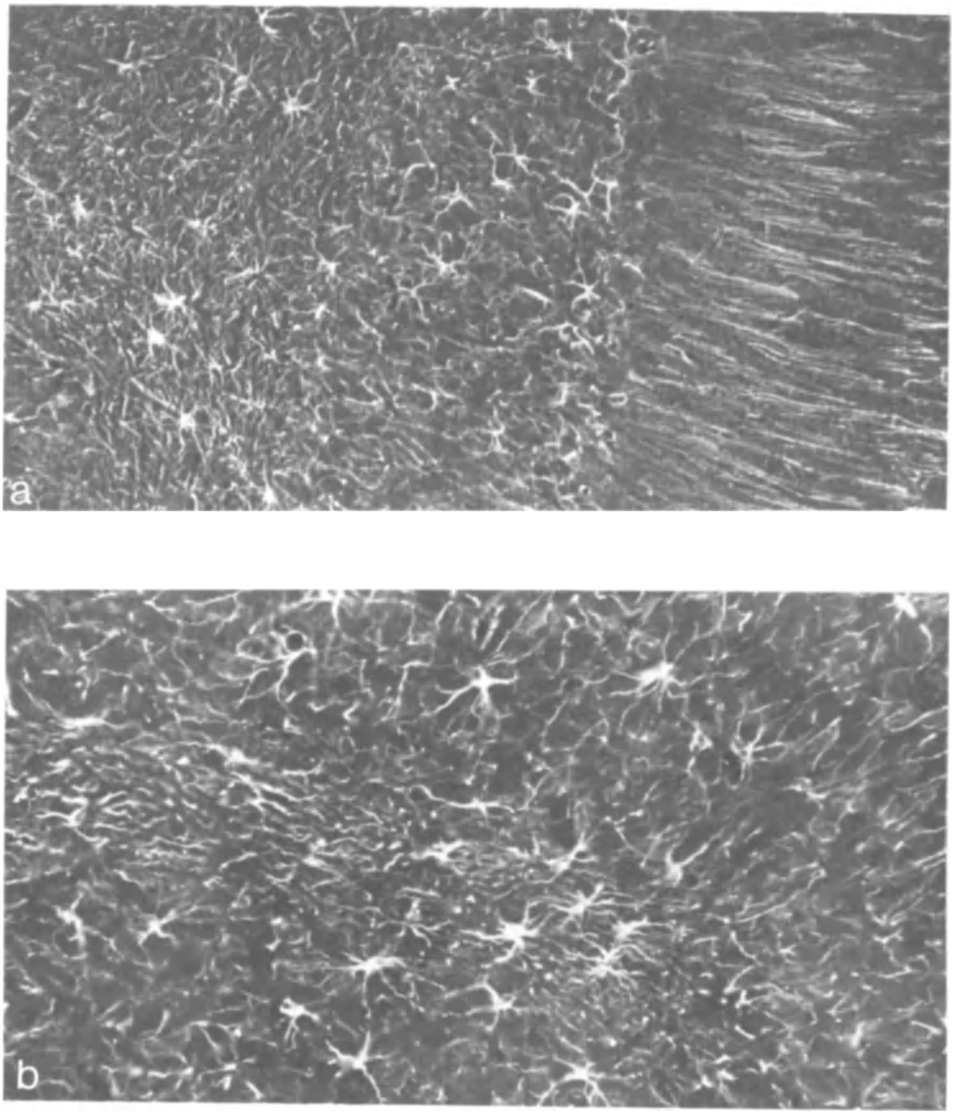


Fig. 2: Fluorescence micrograph of posterior vermis of cerebellum from rat exposed to 14 kV/m, 60 Hz for 4 wks. An increased number of large, multipolar astrocytes(arrow) extend from the Purkinje nerve cells(P)through the granular cell layer into the white matter(w)as seen in low magnification (2a) and at higher magnification (2b).

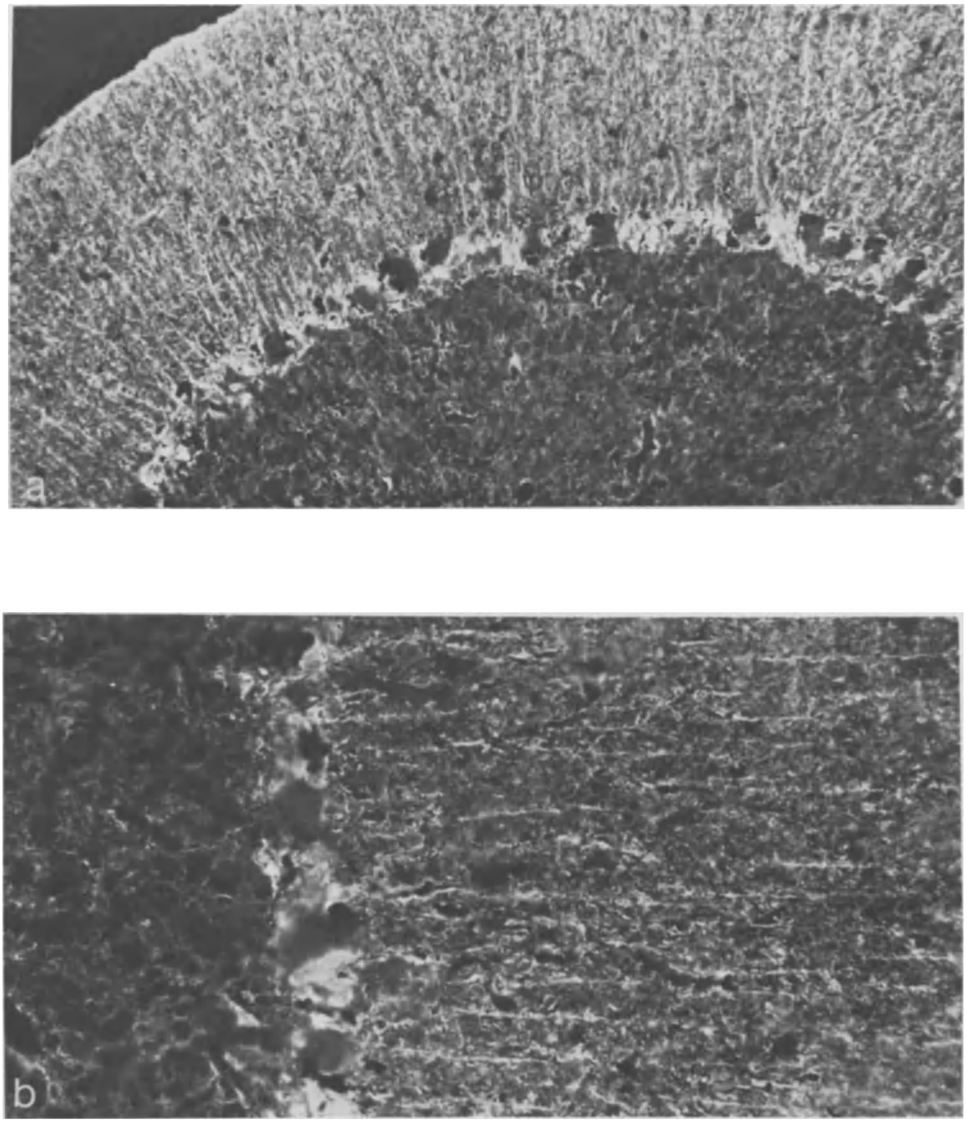


Fig. 3: Fluorescence micrograph of posterior cerebellar vermis of rat exposed for 4 wks from birth to 14 kV/m, 60 Hz processed for demonstration of the calcium-binding protein s-100. Intense staining of increased number of Bergmann radial glial cells(arrow)among the Purkinje nerve cells(P), looking like dark holes. 3b; Higher magnification shows more detail of the glial processes to the rt. and the granular cell layer to the lt.

embedded specimens confirmed the glial reactive changes described. Interestingly enough, the Purkinje nerve cells turned out to contain extremely few lamellar bodies, so frequent in corresponding rabbit cells. However, cytoskeletal alterations as reflected by the distribution and number of filaments and microtubules, were observed as were changes in exposed rabbits.

MICE

Albino mice were exposed to electric fields (60 Hz, 9.44 kV/m) from birth up to the age of four weeks. Their weight gain and behaviour were checked to see whether the exposure induced any differences as compared to controls. Thereafter, specimens from different parts of the nervous system including the cerebellum were examined as for the rabbits and rats.

For mice, immunohistochemical analysis of the distribution of the S-100 protein turned out to be the best marker of reactive glial changes in the cerebellum. The Bergmann glial cells increased in number and size. Furthermore, they showed the same conspicuous change in distribution of S-100 immunoreactive material as previously described, i.e. the normal cytoplasmic diffuse distribution was replaced by enrichment in the nucleus. The astrocytes in the cerebellum of the mouse display only faint immunoreactivity for the GFA protein, which thus turned out not to be as suitable as a second marker of glial reactive changes.

Morphological evaluation of the glial cells using light and electron microscopic methods confirmed the presence of glial changes. However, in mice as in rats the lamellar bodies did not enable proper identification of exposed animals due to their very low frequency. Studies are in progress to evaluate morphometrically the cytoskeletal changes observed in especially Purkinje nerve cells as well as pyramidal nerve cells in hippocampus.

PIGS

Hanford mini pigs were exposed for months to electric fields (60 Hz, about 30 kV/m, 21 hrs per day) at the Batelle Northwest Pacific Laboratories in Richland, Wa. Specimens of brain tissue fixed by immersion became available for examination from three exposed pigs and one control. Preliminary results from this pilot study indicate that glial reactive changes similar to those described above were induced. There is a marked increase in the number of GFA immunoreactive astrocytes in the vermis posterior of cerebellum (Fig. 4). The Purkinje nerve cells are encircled by numerous astrocytes, i.e. Bergmann glial cells (Fig. 4b). Furthermore, the density of the glial meshwork in the granular cell layer is increased (Fig. 4a). The S-100 immunoreactivity is fairly low and therefore the few specimens available does not allow firm statement about possible differences. The electron microscopic studies were impaired by the artefacts induced by the fixation procedure used.

Thus, the limited number of specimens available seems to indicate effects on the nervous system by long term exposure to electric fields. We are continuing with studies on specimens from more pigs, exposed and control, in order to be able to make firm statements of the significance of the glial reactions observed.

COMMENTS

A consistent finding in the several studies reported was the occurrence of reactive glial changes in both the gray and white matter of the cerebellum. There was not only an increased number of glial cells but also increased size of the meshwork of processes observed especially in the cortical layers. However, considerable variations were noticed reflecting species differences as well as, to a certain extent, regional variations. It is likely that these changes are due to disturbed interaction between glial cells and nerve cells. However, we do presently not know whether the glial changes are the primary ones or secondary. Further studies are in progress to elucidate this important question.

Recently, Wilson et al. (1982) reported significant effects on the melatonin hormone secretion in rats after long term exposure to electric fields. Interestingly enough, there does not seem to be any difference in the responses to electric field intensities of 65 kV/m vs 1.2 to 1.3 kV/m. The exposure time factor was more important as significant effects were observed only after three weeks. In the present study we observed very similar glial changes independent of species, frequency (50 or 60 Hz) or intensity. The shape and size of the bodies of rabbits, Hanford mini pigs, rats and mice are very different and ought to result in considerable difference in field enhancements (Kaune et al., 1981). The currents

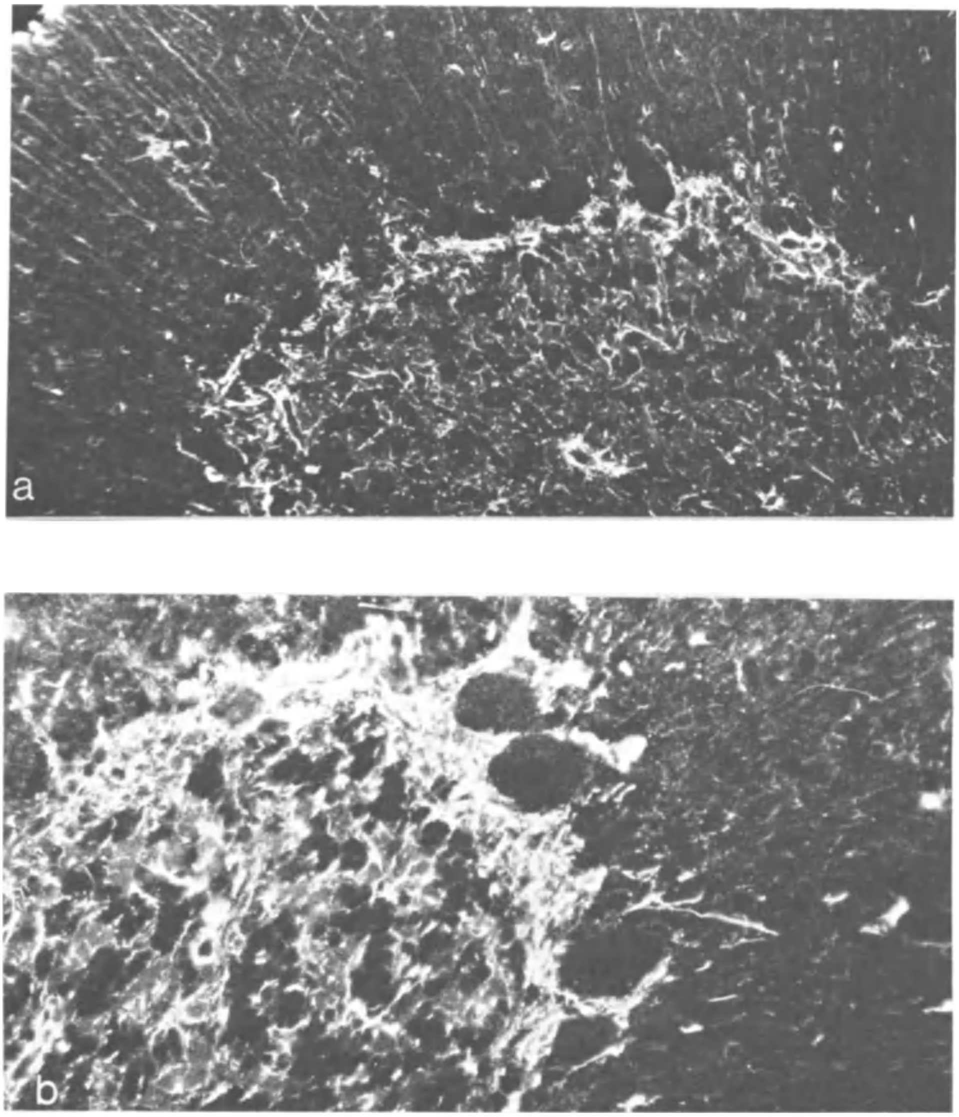


Fig. 4: Fluorescence micrograph of posterior cerebellar vermis of a Hanford mini pig exposed long term to about 30 kV/m, 60 Hz, processed for demonstration of GFA. Glial cell bodies and processes are seen in strongly increased frequency along the Purkinje cell layer(P) and form a network in the granular cell layer(below, rt.) being more extensive than in the control. This is shown in higher magnification in 4b, in which the encircling of Purkinje nerve cells by astrocytes is obvious.

both at the surface of the animals and within their bodies are very different. These results indicate that the exposure schedule and time factors play key roles with regard to induction of effects on nervous tissue.

Our present studies are mainly concentrated upon the study of the interaction between the nerve and glial cells as well as the possible role of the blood vessels. Immunochemical, biochemical and pharmacological techniques will be used in addition to the morphological and immun histochemical ones. Furthermore, extensive morphometric evaluation of the structural changes observed are in progress, hopefully enabling us to evaluate the significance of the changes observed and possible health hazards.

ACKNOWLEDGEMENTS

This study was made possible thanks to collaborative efforts by R.W. Adey, L. Anderson, S. Dyberg, E. Edvardsson, K.G. Haglid, W. Kaune, A.A. Marino, B. Nordström, L.E. Paulsson, R.D. Phillips, L. Rosengren, B. Rozell, P. Sagan, A. Sheppard, and U. Svedin.

REFERENCES

Adey, W.R., 1981, Tissue interactions with nonionizing electromagnetic fields, Physiol. Rev., 61:435.

Becker, R.O. and Marino, A.A., 1982, "Electromagnetism and Life", State University of New York Press, Albany.

Hansson, H.-A., 1981a, Lamellar bodies of Purkinje nerve cells experimentally induced by electric field, Brain Res., 3:47.

Hansson, H.-A., 1981b, Purkinje nerve cell changes caused by electric fields: Ultrastructural studies on long-term effects on rabbits, Med. Biology, 59:103.

Kaune, W.T. and Phillips, R.D., 1980, Comparison of the coupling of grounded humans, swine, and rats to vertical, 60-Hz electric fields, Bioelectromagnetics, 1:117.

Palay, L.S. and Chan-Palay, V., 1974, "Cerebellum", Springer Verlag, Berlin.

Sheppard, A.R. and Eisenbud, M., 1977, "Biological Effects of Electric and Magnetic Fields of Extremely Low Frequency", New York University Press, New York.

Sheppard, A.R., 1983, "Biological Effects of High Voltage as Transmission Lines", Montana Department of Natural Resources and Conservation, Helena.

Wilson, B.W., Anderson, L.E., Hilton, D.I., and Phillips, R.D., 1982, Chronic exposure to 60-Hz electric fields: Effects on pineal function in the rat, Bioelectromagnetics, (in press).

NEUROTRANSMITTER RELEASE STIMULATED IN A CLONAL NERVE CELL LINE BY LOW INTENSITY PULSED MAGNETIC FIELDS

G. Rein and R. Dixey

Department of Medical Electronics
St. Bartholomew's Hospital West Smithfield
London, ENGLAND

Because of our growing interest in the biological activity of electromagnetic fields (EMF), we have set out to establish a model system for studying their mechanism of action at the molecular level. Although *in vitro* preparations of bone (1,2) and nerve cells (3-6) have been used, the populations are heterogeneous and direct effects on specific membrane events are difficult to monitor. Isolation of specific cell types has been used for red blood cells (7). Separation of neuronal cells, however, requires elaborate time-consuming techniques which are likely to affect the integrity of the membranes in the final preparation. Single cell clonal lines of neural crest tumors in culture are likely to be useful, since they express many properties associated with normal mature adrenal chromaffin cells (8-13). Neuroblastoma cells are an example of one such line which is being used for EMF studies (14). Recently, Greene and Tischler developed a more differentiated clonal line, PC12, derived from a rat phaeochromocytoma (12) which also exhibits properties of sympathetic neurones including calcium dependent release of neurotransmitters (8,9,11).

We have previously reported (24) that a 500 Hz signal applied to a helmholz coil arrangement stimulates (^{3}H)-noradrenaline release from these cells, and here summarize those findings and discuss the limitations as well as implications of this experimental approach.

A magnetic field was generated by passing current through a pair of concentric 400 turn coils wired in parallel. These were separated by 14.7 cm, and the experimental dishes placed symmetrically about their midline, at 4.97 and 9.75 cm. The waveform of the applied signal is shown in Figure 1. This waveform was the result of using a pulse generator and driving circuit similar to that used in clinical trials for the treatment of non-union fractures in Newcastle and London (15). The generator gives a train of square pulses with a range of different pulse durations, frequencies and intensities. The changing magnetic field induces currents in the dishes, which decrease to zero at the center. The range of currents so induced was consequently limited to within 20% of the mean by glueing a smaller dish in the center of the larger (60 mm o.d.) petri dish, and plating cells in the annular channel so formed (16).

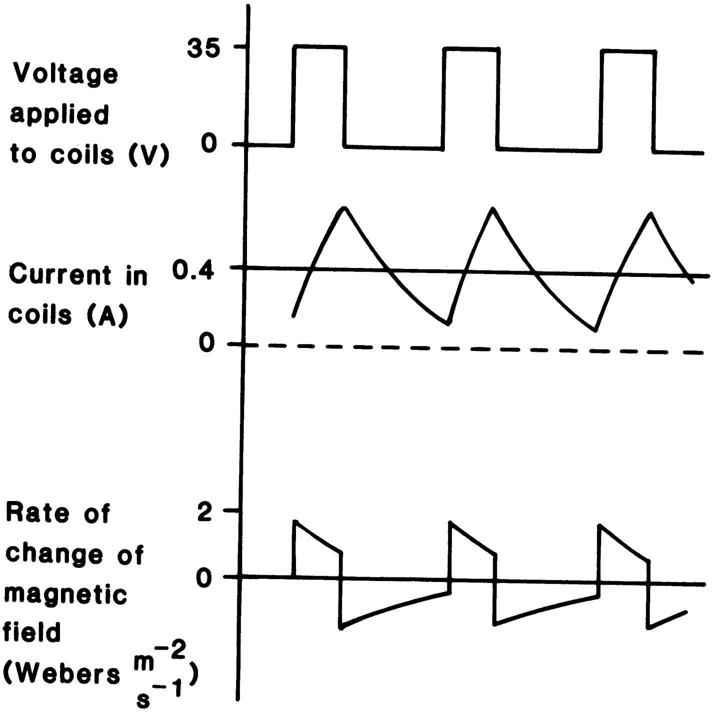


FIGURE 1

For these experiments we used a pulse width of 0.6 ms, with 1.4 ms between each pulse, giving an overall frequency of 500 Hz. The current waveform was smoothed due to the large inductance of the coils. The applied magnetic field will have the same form as the current in the coils and can be calculated from the geometry of the coils and the current (17). The maximum field applied to the sample was 8.5 x 10^{-4} Webers m^{-2} (8.5 gauss) and the minimum 1.6 x 10^{-4} Webers m^{-2} (1.6 gauss). The rate of change of the magnetic field, calculated from the voltage induced in a search coil, varied between +1.8 and -0.9 Webers m^{-2} s^{-1}, and the calculated circulating eddy current density varied between +0.025 amps m^{-2} and -0.013 amps m^{-2}, inducing an electric field ranging from +3.8 x 10^{-2} V/m to -1.9 x 10^{-2} V/m.

PC12 cells were maintained in culture using DMEM medium supplemented with Donner calf and horse serum. We measured (^{3}H)-noradrenaline release by using logarithmic phase cells, plated on collagen, through a modification of Greene and Rein (9). After a 2-hour uptake period, the cells were washed at 30 minute intervals for 3 1/2 hours to deplete cytoplasmic (^{3}H)-noradrenaline stores. A washout curve was enumerated (Figure 2) by changing the medium at 15-minute intervals and measuring the radioactivity of the noradrenaline in the extracellular medium. In order to avoid high values due to cells dislodging during the medium changes, all samples were centrifuged at 2200g for 5 minutes prior to the assay for radioactivity. The pellets were solubilized and counted separately. At predetermined 105 and 120 minute time periods the cells were placed between the coils, and exposed to the magnetic field for 13 minutes. At all other times the cells were placed in a mumetal box to shield from stray magnetic fields present in the incubator.

In all 10 trials where the EMF was applied, release values counts per minute (cpm) obtained at 105 and 120 minutes were always higher than the values obtained from the two preceding time periods. In order to estimate the extent of the observed stimulation, the first six release values for all 22 trials were fitted to the polynomial function $\log Y = \emptyset(x) + N + T$ from which the control curve of Figure 2 is obtained. Y represents the cpm in the extracellular medium, x is the time in minutes, N is a normalizing term allowing direct comparison between dishes, and T is a term inserted for the experimental period (18). In the control dishes ($n = 12$) the observed values for the experimental periods fitted within one standard deviation of those obtained by extrapolating the polynomial function, showing that the washout curve varies exponentially with time. The experimental values (during EMF stimulation) were elevated above the controls with the treatment effect being highly significant ($p < 0.001$).

A percentage stimulation ($\pm$S.E.M.) of 27.5% $\pm$4.9 is obtained for both experimental time periods. Expressing the stimulated release as the difference between the control and experimental values relative to the control value for a given time period. This elevation is likely to be due to increased release of noradrenaline from the cells in monolayer and not due to contamination by cells caused to float by the EMF, since the pellet counts for the experimental periods are not significantly different ($p > 0.5$) from those before or after it (see Table 1). The temperature of the coils was continuously monitored and did not rise above the 37°C of the CO_2 incubator.

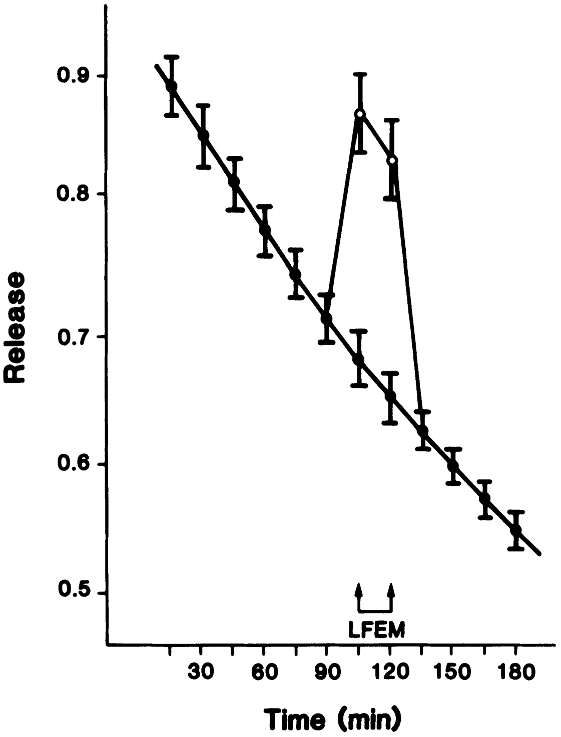


FIGURE 2

Effect of a 500 Hz pulsed magnetic field on noradrenaline release from PC12 cells by regression analysis from a population of 22 curves. All manipulations were carried out in a 37°C warm room. Logarithmic cells (1×10^6 cells per dish) were incubated for 2 h in Dulbecco's modified Eagle's medium containing 2 μCi ^{3}H-(-)NA-(11Ci/mmol) in 7% CO_2 atmosphere at 37°C. The radioactive medium was then removed and the cells washed three times with fresh medium (at 37°C). The cells were incubated for 3.5 h in 1.5 ml fresh medium, with changes after 2 h and then at 30 min intervals. Thereafter, it was changed every 15 min for 12 consecutive periods. The medium from these changes was centrifuged at 2200 g for 5 min. Supernatants were assayed for radioactivity. After each experiment, radioactivity remaining in the cells was assayed by scraping the cells sequentially in 0.5 ml 1 M HCl, 0.5 ml 1 M NaOH and 0.5 ml water. Release values were calculated from the ratio of the c.p.m. in the supernatant, at a given time period, to the total residual intracellular radioactivity at the end of the experiment. The mean total intracellular count (± s.e.m.) was 389,018 ± 26,827. We have shown previously that 84% of the radioactivity released from the cells is unmetabolized NA[12]. ● , Control curve. ○ , Stimulated points.

In eight further trials, 15 mM Mg^{++} was included in the medium for the experimental periods. This procedure did not alter the rate of release in control washout curves, but abolished the stimulated release as shown in Table 1. The stimulated release values in the presence of elevated Mg++ were not significantly different from the controls ($p > 0.3$).

In three of the trials, the cells were bathed in 51 mM K^+ (8) following the period of stimulation by the EMF. In response to this stimulus, the cells released a substantial proportion of their remaining noradrenaline stores, indicating that the cells were still capable of responding to conventional depolarizing stimuli.

DISCUSSION

Our chief objectives in setting up a model system were: (a) to obtain ready access to a stable and homogeneous biological preparation; (b) to study a well defined characteristic of the preparation; (c) to quantify applied field strengths and current intensities precisely, through a simplified sample geometry.

Neurotransmitter release in clonal cell lines met these objectives. The results of our preliminary observations suggest that PC12 cells may be a useful model system for studying the effects of EMF on neurotransmitter function. However, there are certain problems inherent in the model system. First, PC12 is a highly differentiated cell line and consequently requires a specific tissue culture environment to maintain a stable population. One critical factor is the type and batch of sera used, at least with regard to sera available in the U.K. In our hands the responsiveness of PC12 cells to magnetic fields seems to be dependent on the specific batch of sera used. Different batches of sera contain different phospholipids with varying fatty acid compositions which become incorporated into the plasma membrane. A recent report (19) linking membrane fatty acid profiles with depolarization-dependent exocytosis in PC12 cells may provide a key to this problem. Resolution awaits a further study. A second problem is that neurotransmitter release from adrenergic neurones is generally highly temperaturedependent (20), a finding which we have confirmed in our PC12 cells. Although the possibility of large increases in temperature can be effectively ruled out in our system, the elimination of small local increases may be extremely critical.

In general, however, the study of cell lines in culture offers great scope for the study of electromagnetic interactions. Recent studies in these laboratories (21) have shown that low frequency EMF can also modulate the responsiveness of PC12 to nicotinic cholinergic stimulation, and these findings, coupled with the inhibition of the EMF-induced release by elevated Mg^{++} mentioned in

TABLE 1

c.p.m. expressed as % of cell content		TIME (min)			
	n	90	105	120	135
Control	12	0.717±.039	0.686±.046	0.657±.044	0.628±.033
Experimental	10	0.717±.039	0.875±.064	0.838±.061	0.628±.033
Experimental + 15 mM Mg^{++}	8	0.717±0.39	0.637±.015	0.686±.02	0.628±.033
Pellets	10	0.049±.012	0.05 ±.014	0.045±.013	0.06 ±.024

^{3}H-Noradrenaline release from PC12 cells. Release values were generated by regression analysis and expressed graphically as in Figure 2. In the elevated Mg^{++} trials the medium for the 105 and 120 min time periods, during field stimulation, contained 15 mM MgSO$_4$. the aspirated medium was treated as previously described in Figure 2. Pellets obtained from spinning aspirated medium (experimental series) at 2200 x g for 5 min were solubilized in 1.5 ml of 1N NaOH and assayed for radioactivity in 10 ml Instagel. All values are expressed as a percentage of the residual intracellular radioactivity and are shown as mean ± s.e.m.

our previous report, confirm our conclusion that the fields interact directly with the Ca^{++}. dependent exocytosis mechanism present in these cells (8,9,11).

The electric fields induced in this system are of the order of 2×10^{-2} V/m or less. In contrast our PC12 cells display resting potentials of 15 - 20 mV across their plasma membranes (13,22). Since this corresponds to a fixed electric field of some 10^{5} V/m, it seems extremely unlikely that the induced EMFs directly impose electrical charges on the cells' plasma membranes. A more plausible hypothesis is that co-operative interactions are responsible for modulating the release process in these cells (23). The possibility that Ca^{++} binding might be directly affected has been proposed for other systems (4,5).

Release of catecholamines from PC12 can be induced by either depolarizing the cell membrane with elevated K^{+} (8,11), or sodium ionophores (8), or by activating the acetylcholine receptors with nicotine or other cholinergic agonists (9). Release under these conditions is large, rising to 300% over basal rates for maximal concentrations of nicotine, whereas the stimulation we have found is of the order of 20-30%. It is likely that this effect is not optimal. Determination of the limits awaits further studies which vary the EMF intensity, pulse frequency and shape. The difficulties we are experiencing in stabilizing our preparation sufficiently to resolve these small interactive effects are the main obstacles to further progress of this model at the present time.

REFERENCES

1. R.A. Luben, C.D. Cain, M.C.-Y. Chen, D.M. Rosen and W.R. Adey, Proc. Natl. Acad. Sci., 79:4180-4184 (1982).
2. C.W. Archer and N.A. Ratcliffe, J. Exptl. Zoology 225: 243-256 (1983).
3. B.F. Sisken and J.F. Lafferty, J.F. Bioelect. Bioenergetics 5:459-464 (1979).
4. S.M. Bawin and W.R. Adey, Proc. Natl. Acad. Sci., USA 73:1999-2003 (1976).
5. W.R. Adey, Physiol. Rev. 61:435-514, (1981).
6. C.F. Blackman, S.A. Elder, C.M. Weil, S.G. Benane, D.C. Eichinger and D.E. House, Radio Science 14:93-98 (1979).
7. A. Chiabrera, M. Hinsenkamp, A.A. Pilla, J. Ryaby, D. Ponta, A. Belmont, F. Beltrame, M. Grattarola and C. Nicolini, J. Histochem. and Cytochem. 27:375-381 (1979).
8. L.A. Greene and G. Rein, Brain Res. 129:247-263 (1977).
9. L.A. Greene and G. Rein, Brain Res. 138:521-528 (1977).
10. L.A. Greene and G. Rein, Nature 268:349-351 (1977).
11. A.K. Richie, J. Physiol. (London) 286:541-561 (1979).
12. L.A. Greene and A.S. Tischler, Proc. Natl. Acad. Sci. USA 73:2424-2428 (1976).

13. M.A. Dichter, A.S. Tischler and L.A. Greene, *Nature* *268*: 501-504 (1977).
14. S.K. Dutta, B. Ghosh, A. Subramonian and R. Parshad, *Bioelectromagnetics* (in press).
15. B. Kaler, P.O. Byrne, D.D. Milner, *Proc.* *H.P.A.*, Sheffield, England (October 24, 1980).
16. M.J. Lunt and B.W. Watson, *Clin. Phys. Physiol. Measurement* 3:221-225 (1982).
17. M.J. Lunt, *Med. Biol. Eng. Comput.* *24*:501-511 (1982).
18. P. Armitage, *Stat. Methods in Medical Res.*, Blackwell Scientific Publications, Oxford (1971).
19. T.P. Williams and R. McGee, *J. Biol. Chem.* *257*:3491-3500 (1982).
20. D.M. Paton, in: "Mechanisms of Neuronal Transport," D.M. Paton (ed.), Raven Press, New York (1976).
21. R. Dixey (unpublished observations).
22. R. Dixey and G. Korenman (unpublished observations).
23. S.M. Bawin, A. Sheppard and W.R. Adey, *Bioelectrochemistry and Bioenergetics* *3*:67-76 (1978).
24. R. Dixey and G. Rein, *Nature* *296*:253-256 (1982).

NATURAL AC ELECTRIC FIELDS IN AND ABOUT CELLS

Herbert A. Pohl

Department of Physics
Oklahoma State University
Stillwater, OK 74078

ABSTRACT

Living cells are the source of natural ac electrical oscillations, as is shown by several experimental approaches. These include dielectrophoresis, cellular spin resonance, and direct detection. This paper reviews and compares these recent results. The cause and consequences of these natural resonances require clarification.

The natural cellular resonances have been detected in a wide spectrum of cell types, ranging from the primitive bacteria to mammalian cells. They are therefore probably "universal". In one species where they can be readily compared during various phases of the cell life cycle, as in the yeast, Saccharomyces cerevisiae, the electrical oscillations appear to be a maximum at or near mitosis. This brings up questions as to why, how, and when such oscillations would arise.

To the question, "Do these natural oscillations reflect cellular operations which are necessary or are they a frill?"; one is led to assume that they reflect needed processes, for they have presisted through the long evolutionary path. If, for example, they reflect events necessary in the cell's reproductive sequence, this implies that there may be an electrical aspect to cellular growth and its control. How this might be related to the phenomnon of "contact" or "density" inhibition of cell growth is discussed. Several mechanisms whereby such natural ac oscillations may arise in cells are suggested and discussed.

INTRODUCTION

Active search for natural electrical oscillations and ac fields in living cells began about 15 years ago, following the initiative of Herbert Fröhlich (1968,1980). He predicted that there would be very high frequency oscillations (ca.100 GHz), but left open the possibility that cooperative oscillations of a rather lower frequency could also be expected. We shall focus in the present paper upon low frequency electrical oscillations; for the evidence concerning the very high frequency range has been adduced and summarized by Fröhlich (1980, and cf. this volume). More recent evidence from laser Raman spectroscopy (Webb, 1980, and cf. this volume) indicates that metabolically active cells have responses in the multigigaherz range. It has been suggested (Del Guidice et al., 1982) that the Fröhlich mechanism, that of phonon interactions with loosely coupled dipolar systems acting cooperatively in the cell, is also augmented by solitons (Davydov, 1980) so as to produce low frequency oscillations. Much remains to be learned.

To assist the reader in assessing the experimental data, it may be helpful to have in mind several likely mechanisms for the observations to be considered. Let us focus on the likely causes for those frequencies lower than about 100 MHz. The natural electrical cellular oscillations in this lower range can be considered to arise by several mechanisms. In considering any particular mechanism we must bear in mind that field effects from the cell, and acting upon external systems, must be considered as due to electrical distortions at or near the cellular surface, for the high conductivity of the cell interior can be expected to mask or severely damp low frequency fields arising from deep within the cell.

This broad restriction applies especially to detection techniques such as micro-dielectrophoresis (see below), cellular spin resonance, or direct detection methods. Moreover, the effects of a cellular ac field upon external objects such as the test particles used in micro-dielectrophoresis can also be expected to be of very short range unless the conductivity (i.e. the screening length) of the surrounding medium is very low. To say it in another way, one should expect that only surface-associated and assymetric charge distributions of the cell would be detected by nearby test objects. The effects of natural ac electrical oscillations can be expected to be short-ranged and quite ineffective unless sought for in media of low conductivity. Since we are considering the natural medium, water, the water must therefore be relatively free of ions if our search for natural cellular ac oscillations is to be successful. The outer extent of the natural ac oscillations will be higher at the higher frequencies.

Among the mechanisms which should be considered to be involved in natural ac oscillations are those involving the Fröhlich mechanism of cooperative dipolar oscillations driven by cellular metabolic energy, and mechanisms involving oscillating chemical reactions capable of producing charge waves. The latter could well involve limit cycle operation as discussed by Fröhlich (1977), by Kaiser (1982, and cf. this volume), and by Albanese (1983, this volume). A model for oscillating chemical reactions occuring in a cell, and giving rise to charge waves was discussed by Pohl and Braden (1982).

In this model it is proposed that the natural intrinsic ac oscillating cellular dipoles stem from oscillating chemical reactions that couple spatially within the cell to mobile sets of ions so as to produce charge density waves. Based upon what is known of several well-studied oscillating reactions, (Noyes and Field, 1974; Noyes, 1977; Eppstein et al., 1981), the cyclic course of the reaction runs between an ionic and a (branching) free radical phase. As the reaction develops, it develops in both temporal and spatial coordinates from the initiating site(s). It is therefore clear that if, during the ionic phase of the reaction cycle, the outward speed of the positive ions does not exactly match that of the negatives, then a charge wave will develop. It is an example of a longitudinal electrical wave. Such charge waves, if uncoordinated, could begin at any time and at any place within the cell, and be incoherent. However, as these charge waves develop, they encounter structures within the cell which will tend to make them run in parallel. A collimation of the charge waves will follow. Since they interact electrically, a Coulombic correlation (in the sense of a Bose-Einstein condensation) will tend to be formed. A coherent set of charge waves will thus involve large regions of the cell in a cooperative oscillating dipole or multipole. Such a testable model suggests a variety of experimental studies as to the energy source, the magnitude, the control, and causes.

Oscillating reactions in biological systems are numerous and well-known. Excellent reviews are to be had on the subject. (Treherne et al., 1979; Berridge and Rapp, 1979; Rapp, 1979.) A number of well-known structures within cells are candidates for the means by which such cyclic charge waves might be channelled and made more or less parallel. These include the walls of the endoplasmic reticulum, the laminae of the cristae in mitochondria or of the granae in chloroplasts, or the mitotic spindle apparatus.

The charge waves as proposed above could travel geometrically parallel to, or at right angles to the laminar structures restricting them in the cell. Kinetic data on the presumed charge waves imply that the charge waves travel across rather than along the laminae, especially in the case of high frequency oscillations. (Pohl and Braden, 1982) A theoretical analysis by Schmidt and

Ortoleva (1979) suggests how oscillating reactions might be coupled to electrical charge waves.

For the natural cellular oscillations in the frequency range of 10 Hz to 1 MHz, it is likely that the ionic double layers at the many interfaces within the cell would play a role as a source of the set of mobile ions involved. Conversely, it is also possible that mechanical motion acting on biological objects produces electromagnetic waves having feedback into the natural ac oscillations of cells. We refer here to the recently discovered mechanically-induced electromagnetic radiation (M.I.S.E.R.) mechanism. (Pohl, 1980a).

Another possible mechanism capable of producing resonant oscillations from cells is that which would be produced by local fluctuations in the charge density at the surface of the cell, which arise in turn from internal reaction oscillations. For example, let there be a local charge density fluctuation dq, which occurs at a given point at the cell surface at time t = 0. It will then affect the nearby charge distribution in the ionic double layer at the cell-medium interface surrounding the cell. This perturbation will then travel around the cell periphery and reverberate back to the origin with a frequency f, in the order of

$$f = v/1, \qquad (1)$$

where v is the velocity of the disturbance, and 1 is the half-perimeter. A rough estimate of the frequency of this reverberation can be had by assuming that the velocity is in the order of that of sound, and 1 = (pi)x (cell radius). For a cell of radius 10 microns, the frequency will be in the order of

$$f = v/1 = (1500\ \mathrm{m/s})/(3.14 \times 10\ \mathrm{microns}) = 50\ \mathrm{MHz}.$$

It is felt that more realistic values of the velocity of propagation, v, of the perturbation of the ionic double layer might lead to somewhat lower values of the reverberation frequency involved in this type of electrical oscillation about cells. The ionic double layer has been considered, for example, to be comprised of a tightly held and of a loosely bound set of counterions. (Dukhin & Shilov, 1974; Pohl, 1978) The more loosely bound counterion set will be expected to reverberate at a much lower frequency, probably lower by a factor of about 1000, judging from the results of Dukhin and Shilov. The intensity of this mode will depend upon the feedback between the internal and external waves. Consideration of these details will doubtless lead to clarification of the model.

To sum up briefly, there are several reasonable mechanisms whereby natural ac electrical oscillations could arise in cells. In viewing the experimental results one can picture the electrical oscillations:

(a) To arise from cooperative dipolar oscillations fed by the metabolic energy. (Fröhlich, 1968, 1980) This mechanism can account well for oscillations at very high frequencies (multigigaherz range), and be expected to have low frequency modes in certain cases (soft modes).

(b) To arise from charge waves resulting from the presence of oscillating chemical reactions collimated by internal cellular structures, and made coherent by coulombic interactions. Because of the high conductivity and resultant electrical screening at low frequencies, only the features of these charge waves active near the cellular surface can be expected to be strongly felt outside the cell.

(c) To arise from charge wave reverberations in the exterior ionic double layers associated with the membrane or with the cell wall (or "buffy coat"), with these charge waves triggered or originating from aperiodic or periodic perturbations from within the cell.

THE EXPERIMENTAL EVIDENCE

There are several lines of experiment which provide evidence for the existence of natural ac electrical oscillations in living cells. Rowlands et al. (1982a, 1982b, 1983) showed that interactions of extraordinarily long range exist between living red blood cells in their plasma. Their measurements imply that the interactions extend out to some 4 microns from the cell periphery, a fact that strongly suggests the presence of long range cooperative interactions of the type proposed by Fröhlich.

Several experimental techniques have been used to demonstrate the presence of natural electrical ac oscillations in the frequency range below 100 GHz. There are three such techniques (1) dielectrophoresis (DEP), (2) cellular spin resonance (CSR), and (3) direct instrumental detection. In this section we briefly review the recent evidence from these.

Dielectrophoresis (DEP, for short) is the motion of bulk or particulate matter induced by the action of nonuniform electric fields (Pohl, 1951, 1978). It can be used to explore for the presence of nonuniform fields about cells or other objects. In the present instance the motion of tiny polarizable particles near cells is used to help decide if ac fields are present near the cell surface. The motion will be towards the region of highest field intensity (positive dielectrophoresis) if the effective dielectric constant of the tiny polarizable particle exceeds that of the suspending medium. Contrarywise, the motion of the particles will be away from the region of greatest field intensity (repulsion, or negative dielectrophoresis) if they have a lower effective dielectric constant than that of the medium. Because the phenomenon occurs on a microscopic scale, it is referred to as micro-dielectrophoresis (micro-DEP).

Micro-DEP experiments can be done with simple apparatus, since no external fields, other than of the cells themselves are required. In brief, one puts a mixture of the cells and selected tiny (about 2 micron diameter) particles under the microscope and counts the fraction of particles associated with the cells. The average number, n, of particles per cell, and the concentration of particles, p, is taken. The ratio, n/p, reflects the degree of association of the particles with the cell. The n/p ratio for particles of high polarizability is then compared with that obtained for particles of low polarizability. Examples of useful particles of high polarizability are: $BaTiO_3$, $NaNbO_3$, $LiNbO_3$, or the polyacenequinone radical polymer, made from anthraquinone and pyromellitic dianhydride (Pohl and Wyhof, 1972). The latter is typical of a class of aromatic polymers having very high dielectric constants (1000 to 300,000) attributable to the phenomenon of nomadic polarization (Pohl and Pollack, 1977). These test powder materials all have dielectric constants in the order of 500 or higher. The dielectric constant of water, the major component of the medium is 78. As examples of useful particles of low polarizability we have: SiO_2, Al_2O_3, and $BaSO_4$. These have dielectric constants lower than about 13. This simple experiment requires mainly that the suspending medium be of low conductivity and of the proper osmotic character so that the ac fields can be adequately transmitted without undue absorption in the medium. A specific conductivity lower than about 10^{-5} mho/cm is usually sufficient for frequencies in the range of 10kHz or higher. Static fields are immediately suppressed even in doubly-distilled water to very short ranges, in the order of nanometers. Long range effects upon the particles, such as we are discussing here and observed to be in the order of micrometers, cannot be ascribed to the effects of static electric fields.

The presence of natural ac oscillating fields near cells is confirmed by studies of micro-DEP in several ways. The n/p tests show that there is a preference for particles of high polarizability over those of low polarizability:

A) Which is shown by living, not by dead cells (Pohl, 1981b).

B) That is supressable by metabolic inhibitors such as would block ATPases (Pohl et al. 1981).

C) Which is maximal at or near mitosis, as observed in the yeast Saccharomyces cerevisiae . The stages in the life cycle are conveniently determinable by observing the bud development (Pohl et al., 1981).

D) Which is suppressable in accord with expectation from dielectric theory by simply increasing the conductivity of the medium, which in turn raises its effective dielectric constant to exceed that of the test powder and erases, in effect, the DEP forces. (Pohl, 1980a,b,b; 1981b; Pohl et al., 1981). The effective dielectric constant is taken to be the absolute value of the complex dielectric constant,

$$K = K' - jK'' = K' - js/(e_o w).$$

$$K_{eff} = ((K')^2 + (K'')^2)^{1/2}$$

where K is the complex relative dielectric constant, j is the square root of minus one, s is the specific conductivity, K' is the in-phase and K" is the out-of-phase dielectric constant, and e_o is the permittivity of free space, and w is the angular frequency of the electric field. We note from the above equation that the effective dielectric constant, and hence the DEP forces, involves the presence of all charges present ... those seen in dielectric constant measurements and those seen in conduction measurements.

E) Which can be observed for a variety of test powders of widely differing chemical natures (Pohl, 1980a,b,1981b; Pohl et al., 1981, Roy et al., 1981).

F) That is observable in a wide range of organisms, including bacteria, fungi, algae, as well as in avian blood cells, and mammalian fibroblasts of embryonic, normal, and tumor origins. (Pohl, 1980b,c,1981b; Pohl et al. 1981). (See Table I.)

G) Which varies with the conductivity of the medium in a manner that yields an estimate of the minimum frequency of the natural ac fields (about 10 kHz). This agrees with that deduced from CSR experiments (Pohl, 1980c, 1981a).

H) Which provides an estimate of the field strength provided by the cell at its surface (ca. 100 V/cm) consonant with that inferred from torque measurements in the CSR studies (Pohl, 1980c, 1981a). The measurements in this case are made with the ferroelectric powder $BaTiO_3$.

The force exerted by the natural ac fields of cells upon external test objects such as the tiny powder particles is expected to be of very short range. This force varies as Z^{-7}, where Z is the distance between the centers of the cell and the test object. The DEP force, it will be recalled, is proportional to the gradient of the square of the field intensity. Since the field about a dipole varies approximately as Z^{-3}, the gradient of its square will vary as X^{-7}. This implies that the DEP force upon the test particles such as $NaNbO_3$ or SiO_2 will be of very short range about the cells ...in the order of magnitude of a cell diameter.

The natural question arises, "Is the observed preference of living cells for polarizable test particles in the micro-DEP experiments possibly due to static fields?" Several facts rebut this, and give evidence that the preference is an ac phenomenon. First, the evidence cited in points (D) and (G) above show that the presence of even slight conduction in the medium due say, to 100 micromolar NaCl or KCl suffices to suppress the effect, and does so in accord with the theory for dielectrics in ac fields. Second, as is well known from the theory of polar media such as water, the static fields are dampled out by distances of the order of magnitude of a Debye relaxation length. Because this is in the order of only about 100 Angstroms or less in aqueous media, and the micro-DEP effects are observed to be active over distances many-fold larger (several microns), it is therefore quite unlikely that static field effects are dominant here. We conclude then that the evidence from micro-DEP experiments supports the existence of natural ac oscillations by cells.

Let us now turn to the evidence for natural ac oscillations as gotten from the study of cellular spin resonance (CSR). The application of ac fields to a suspension of cells can cause them to spin or rotate selectively. This provides a new and straightforward method for determining the dielectric properties of individual cells or even of inert particles. The technique is particularly useful and transparent when rotating fields, such as those produced by three or four-pole electrode systems, are used. (Mischel, Voss, and Pohl, 1982; Arnold and Zimmermann, 1982a,b; Pohl, 1983). The techniques and the theory are now well in hand, and future progress in this exciting new area should be rapid. An especially interesting aspect of cellular spin in applied ac fields is that in which more or less parallel electrodes are used, and a sinusoidal field is applied.

A number of researchers have observed the spinning of cells while in a simple two-pole ac field. (Teixera-Pinto et al., 1960; Furedi and Valentine, 1962; Pohl and Crane, 1971; Pohl, 1978; Mischel and Lamprecht, 1980; Pohl, 1981a; Zimmermann et al., 1981; Pohl and Braden, 1982). The first account of the resonant spinning of living cells was that of Pohl and Crane (1971). They observed that yeast Saccharomyces cerevisiae, subject to an ac electric field between two parallel wire electrodes, spun in a sharply responsive manner to the frequency of the applied field. Cells were seen to spin alone whether against an electrode or out in the suspension, and to spin collectively while in pearl-chain stacks. The resonant frequency, moreover, depended upon the age of the colony examined. The sharply resonant nature of the spinning response led to the use of the descriptive term, "cellular spin resonance" (CSR). Michel and Lamprecht (1980) quantitatively studied the rate of spin of budding yeast cells. Zimmermann, Vienken, and Pilwat (1981) described the CSR of a number of species, including human erythrocytes and their ghosts, Friend cells, and mesophyll protoplasts of Avena sativa. The variation of the CSR spectrum with phase in the life cycle in yeast cells (Saccharomyces cerevisiae) has been studied by Rivera, (1983).

There appear to be two circumstances for the occurrence of CSR. On the one hand there is the spin of lone cells while little affected by local field perturbations such as those due to rotating fields or delayed polarizations. On the other hand there is the much more general phenomenon of spin while under the influence of rotating ac fields.

The spinning of lone cells in two-pole ac fields was reported a number of workers, including Pohl and Crane (1971), Chen (1973), Pohl and Braden (1982), and by Mischel and Lamprecht (1983). For several physical reasons, the observation of lone cells spinning in the mid region of a two-pole field must be of brief and occasional character, for the cells in that region usually move rapidly towards an electrode or join with other cells to form pearl chains under the influence of DEP. In addition to the action of DEP in causing the lone cells to leave the mid-field region there are other dispersive actions such as thermal upsets, ionic injection, etc causing fluid streaming. Lone cells are, however, readily observed to spin freely and in a resonant way when against a mirror-smooth Pt electrode, as all investigators, but especially Mischel and Lamprecht (1983) have noted. The theory for the CSR of lone cells spinning in CSR in a two-pole ac field was discussed by Pohl (1981a, 1982). At the present writing, only one class of mechanisms clearly accounts for the CSR of lone live cells; i.e., where the frequency of the cellular spin rate is much less than that of the applied field, and where the spinning ceases upon the death of the cell. These mechanisms for the natural cellular dipolar oscillation exist and can interact with the applied field to provide a rotational torque.

It should be commented that the natural cellular oscillations need not merely be simply dipolar, but may oscillate as linear quadripoles or higher multipoles. In view of the possibly weak character of the natural electrical oscillations, it may also be expected that the cellular oscillations would be affected by the externally applied fields such as to "pull" or alter their frequency into resonance with that of the applied field. This would cause the CSR spectrum of cells to consist of peaks somewhat broadened by the frequency pulling, or as the expression goes in dealing with electrical circuitry, by "frequency motor-boating".

Lone cells can be made to spin by static fields, or even in the presence of very low frequency fields through the mechanism of "rotational conduction". In this case the spinning of bodies, even spherical ones, is induced by the action of ionic currents that produce surface-charge dipoles upon the particles having a conductivity or dielectric constant differing from that of the supporting fluid medium. (Pohl, 1978, pp. 135-41). In this case it is expected that the spinning rate will be low. It is also evident that CSR can be evoked even in very high frequency ac fields if natural cellular ac dipoles interact with the externally applied ac fields. The rate of cellular spinning would then be expected to be in the order of magnitude of the difference between the frequencies of the natural dipolar frequency and that of the applied ac field. (Pohl, 1981, 1982)

$$f_{cell} \cong f_{field} - f_{nat\ dipole}$$

The latter mechanism can account for the CSR of lone cells spinning at a low rate (e.g. 0.1 to 10 Hz) while the applied frequency is much higher (e.g. 10 kH to 1 MHz).

While using two-electrode systems to study the CSR of living cells, Pohl and Braden (1982) observed in the case of yeast cells (Saccharomyces cerevisiae) that the CSR spectrum contained only one or two peaks. The cells spun at a rate of about 0.1 to 30 Hz when the applied frequency was in the range of several KHz. The cells spun in a sharply resonant manner with the applied frequency. The CSR was observed to vary with the phase of the life cycle, and with the cell type. Living cells responded readily. Dead cells whether killed by heating or by phenol, exhibited little or no response when observed as lone cells, well removed from the others. Sharp CSR was also observable with live murine sarcoma cells. As a further test of the presence of natural cellular ac oscillations, they compared the CSR of the cells in the presence of pulsed dc and of sinusoidal ac fields. They used both sinusoidal ac and pulsed dc fields. They found the ac and pulsed dc spectra to be similar, and that the cells continued to spin in either field of the same basal frequency and applied field strength. Were cells possessed only of dipoles induced by the external field, they would be expected to librate but not spin when present as lone cells in pulsed two-pole ac fields.

Particles can also be made to spin by the presence of rotational electric fields. Here the rotating field induces a polarization which need not be instantaneous, but lags the field. The interaction of the delayed polarization with rotated field then produces a torque capable of spinning the particle. The particle need not be a living cell, but might be a vesicle, a dead cell, or even an inanimate mineral particle. Rotational fields can be provided by the use of phased three or four pole electrode systems, for example, and this method is useful for determining the dielectric properties as a relaxation spectrum of small particles (Pohl, 1983). Rotational fields capable of causing particle rotation can also arise from the interaction of the particles with each other while present in an otherwise homogeneous field. Consider the case of cells, say, subject to an ac field provided by a two-pole field, but also very near another particle. The field due to the induced dipole of the second particle is phase-delayed compared to that of the applied field, and is therefore capable of providing a rotating field at the first cell, and thereby evoking a rotation. The spinning of cells by this mechanism is easily produced and observed (Holzapfel et al., 1982). The CSR of crowded cells in a two-pole ac field is usually due to this mechanism. The CSR of lone cells in such a two-pole ac field cannot, however be due to such rotational field, and another mechanism such as that due to the interaction of natural cellular oscillating dipoles is needed.

The evidence that living cells naturally produce ac oscillating fields, as derived from CSR studies supports the conclusions reached from the studies of micro-dielectrophoresis. The CSR evidence is not as firm as that for that obtained from micro-DEP studies, in view of the fact that there are several ways to interpret the CSR studies. One must be careful in assigning observable CSR as being due to natural cellular ac oscillations in view of the fact that rotating fields due to the presence of a nearby polarizable object, off-axis from the lines of the applied external field, can evoke spinning. Careful observation to assure that the spinning lone cell is well away from other polarizable objects is required.

Direct instrumental observations of natural ac oscillations of cells have been made by several investigators. The presence of such oscillations in the frequency range of 50 to 500 Hz has been reported from studies of nerve cells (cf. Treherne et al., 1979). Oscillating enzyme reactions have been studied and are well recognized (Hess and Chance, 1978). Action potentials of macrophages derived from human monocytes were observed to oscillate at frequencies of from 20 to about 80 Hz (McCann et al, 1983). The power spectra of mormyrid fish (Pollimyrus isidori) ranges from 1000 to about 50,000 Hz and is involved with the electric organs (Westby and Kirschbaum, 1982). The output is used by such "weakly electric" fish as the South American gymnotid fish and the mormyriform fish of Africa for "electrolocation" in their surroundings. Yano et al. (1981) have directly observed much slower oscillations (at a few Herz) of the syncytiotrophoblast of the human term placenta. Similar observations on murine fibroblasts were made by Okada et al. (1979). Very recent direct instrumental observations of electrical oscillations of high frequency have been reported by Smith (1983). His group used a Hewlett Packard 8553B Spectrum Analyzer and found the yeast, Saccharomyces cerevisiae, to emit oscillations in the neighborhood of 8 MHz. Rivera et al. (1983) recently observed similar results with yeast. The alga, Netrium digitus was observed to provide signals in the order of microvolts at several frequency bands at about 7 and 33 kHz. Detection was done in a shield enclosure with the aid of Pt electrodes externally placed in the medium at the ends of the cell. A preamplifier boosted the signals for input to a wave analyzer (Hewlett Packard, Model 312A) (Fig. 1).

There are, then, three lines of experimental evidence confirming the existence of natural cellular ac oscillations: that from micro-DEP, from CSR, and from direct instrumental observation.

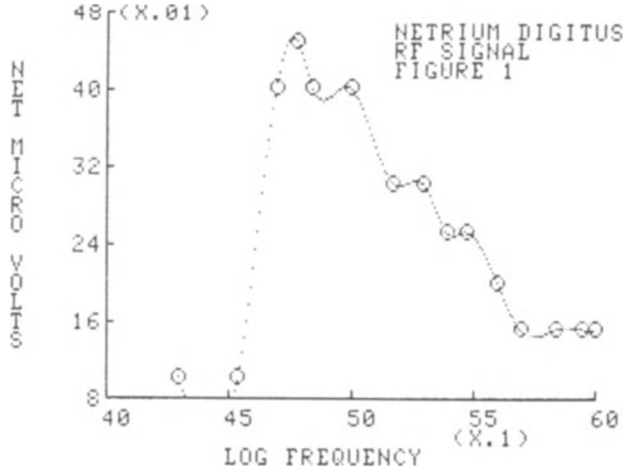


Figure 1

Much more needs to be done to confirm the observations, to establish the nature and range of the oscillations, and to understand their origins and function in the physiology of the cell.

We are presented therefore with the real probability that cells produce natural ac fields. It is fitting then to ask where, how and why such oscillations should arise in living cells. We have, in effect two main experimental results at present. The first broad conclusion from the presently available data is that the cells of many types given evidence of the presence of natural oscillating electric fields. The micro-DEP studies, in which fields about cells are detected by the way in which the cells favor the accumulation of highly polarizable particles over that of weakly polarizable ones, show that the ac fields about cells occur in a wide range of cellular species. This is evident in Table 1.

It is observed from the data of Table 1 that the preference for highly polarizable particles over that for weakly polarizable ones is ubiquitous, and it is a common feature in all of the wide range of cell types.

The second broad conclusion about the presence of natural ac oscillations in cells again comes from micro-DEP studies. The n/p

Table 1. The Preferences of Living Cells for Highly Polarizable Particles ... (as determined with micro-dielectrophoresis).

CELL TYPE	RATIO,POLARIZABLE TO NON-POLARIZABLE PARTICLES;(n/p)polar to (n/p)non-polar.
Bacteria, Bacillus cereus, synchronized	4
Yeast, Saccharomyces cerevisiae	2.0
Alga, Chlorella pyrenoidosa	2.2
Avian, Chicken red blood cell	2.0
Mammalian, Mouse "L" fibroblasts	2.0
Mammalian, Mouse Ascites tumor fibroblasts	2.2
Mammalian, Mouse, Fetal fibroblasts	2.0

values indicating the strength of the preference of the cells for polarizable particles were observed to be maximal at or near mitosis. These results, obtained on the yeast saccharomyces cerevisiae can be easily obtained because the phase in the life cycle is readily determined from the shape of the cell (Hartwell, et al., 1974). The variations of the n/p values for the highly polarizable particles $BaTiO_3$, or $NaNbO_3$ show a peak at the period in the yeast life cycle when the small budding state is attained, i.e., when the cell is at or near mitosis.

Given the evidence that cells do indeed produce natural ac oscillations, and given the two principal results, i.e., that (1) the oscillations are "universal" among the species, and (2) that the oscillations are maximal at or near mitosis; it is interesting to speculate as to their meaning. Perhaps the first question to pose is: "Are the observed natural ac oscillations evidence of a process which is merely an unnecessary frill, or do they reflect a necessary process in the life processes of cells?"; The first principal result, the "universality" of the oscillations, indicates that the natural ac oscillations have been carried on for billions of years throughout the evolutionary process. Nature is not inclined to be wasteful. With this in mind it is a fair presumption that indeed the ac oscillations reflect phenomena which are essential.
It is also likely that the oscillatory phenomena must preceed or accompany the reproductive process, as evidenced by the second principal observation. Together, these suggest that there is an electrical (ac) aspect to the process of reproduction in cells.

We may then make the provisional postulate:
Electrical oscillations must accompany cellular reproduction, and that without them, the reproductive process cannot preceed. This assumption can be tested and examined experimentally in a number of ways. The necessity for electrical oscillations to

accompany the process of cellular reproduction may play a role for example, in the well known process of "contact" or "density" inhibition of cell growth. Contact inhibition is often observable when cells are cultured in the laboratory, for their growth is seen to be checked as the cells become confluent and cover the growth medium.

The usual explanations given for such contact or density inhibition suggest that it occurs because of a shortage of food or, alternatively, because cells put out growth inhibitor factors. But if these were the only explanations, then growth would be expected to resume when fresh medium is flushed across the colony. It does not, however, which implies that there must be some other factor(s) at work. Let us consider an electrical one. If a cell must 'oscillate' to reproduce, as suggested above, and the oscillation is an electrical one that can be damped out or suppressed by the presence of neighboring cells, but not necessarily by the presence of the normal cellular medium, then a simple mechanism can be suggested. One such mechanism which suggests itself is the following: In the spectrum of electrical absorptions characteristic of any material there are normally one or two sharp regions in the spectrum where strong energy absorption occurs. If a feeble electrical oscillator tries to oscillate in the medium at that particular frequency it will be strongly damped. It will only succeed in maintaining its output if it can be permitted to shift to a frequency at which the damping is so small that the energy drain is acceptable. With this picture in mind, let us view the cell as being normally equipped to leave the G_0 and to enter the S-phase with an electrical oscillator system producing its charge waves at a particular frequency, f_1. If this frequency is relatively undamped by the surroundings, the cell will be free to reproduce. This might be the case, for example, of a cell more or less alone in the medium having a strong damping only at a frequency, f_p, far removed from the frequency f_1. But if the cell is amid other cells, and the strong damping frequency of this melange now coincides with the frequency f_1, the cell will then be unable to reproduce, for its necessary oscillations cannot now be maintained. According to this picture, in other words, it will be the density of cells surrounding the individual cell that will modify the electrical neighborhood, and control the ability of the cell to reproduce. Once the local cell density is sufficient to produce high electrical damping around the cell in question, its further oscillation at that frequency is repressed, and the reproductive process also repressed. Such a model can be tested in a number of ways, as by the deliberate injection of foreign material capable of electrically initiating the damping characteristics of massed cells, or by using electrical means to augment the oscillations at the presumed frequency, f_1, etc.

How then are we to account for the well known ability of cells to grow on certain occasions while in a high concentration of other cells? There are at least four such occasions in the life of mammalian cells, for example. These are: during embryonic growth, during normal body cell replacement, during wound healing, and during the growth of tumors. During these periods it will be necessary, in this model, for the cellular oscillations to occur at a frequency somewhat removed from that of the surrounding cell mass would damp out or suppress them. We suggest that this can be accomplished by a shift in the biochemistry of oscillating reaction to a frequency either above or below that of the neighborhood's high loss frequency, and that this frequency shift is under the control of a normally repressed gene. It is interesting in this connection to note that Ahuja and Anders (1977) reported on finding a gene, (Tu gene) which had a repressor gene (R gene) judged to lie on a separate portion of the genetic material, and that these genes were present in all cells capable of reproducing. They associated the control of cell growth with the activity of such a gene pair. It is interesting to ask if it is these genes which control the frequency f_1, and cause it to shift into the frequency range required for the successful operation of the reproductive cycle. Again, these ideas are testable in a variety of ways, and the testing of them may provide interesting insights into the mechanisms of cellular growth control.

ACKNOWLEDGEMENT

This research was supported by the Pohl Cancer Research Laboratory, Inc.

REFERENCES

Ahuja, M. R., and Anders, F., (1977) in RECENT ADVANCES IN CANCER RESEARCH, Vol. I, edited by R. C. Gallo; CRC Press pp. 103-117.

Albanese, R. A. (1983) "Radiofrequency radiation and chemical reaction dynamics," this volume.

Arnold, W. M., and Zimmermann, U. (1982a) Naturwissenschaften 69, 297.

Arnold, W. M., and Zimmermann, U. (1982b) Z. Naturforsch. 37c, 908.

Berridge, M. J. and Rapp, P. E. (1979) J. Exp. Biol. 81, 217.

Chen, C. S. (1973) "On the Nature and Origins of Biological Dielectrophoresis," Ph.D. Thesis, Oklahoma State University, Stillwater, OK 74078.

Davydov, A. S. (1980), Sov. Phys. JETP, 51, 397.

Del Guidice, E., Doglia, S., and Milani, M. (1982) Physica Scripta 26, 232.

Dukhin, S. S., and Shilov, V. N., (1974) DIELECTRIC PHENOMENA AND THE DOUBLE LAYER IN DISPERSE SYSTEMS AND POLYELECTROLYTES, J. Wiley & Sons, N.Y.

Epstein, I. R., Kustin, K., de Kepper, and Orban, M. (1981) J. Amer. Chem. Soc. 103, 2133.
Fröhlich, H., (1968) Int. J. Quantum Chem. 2, 641.
Fröhlich, H., (1977) Neurosci. Res. Program Bull. 15, 67.
Fröhlich, H., (1980) Adv. Electron. Electron Physics 53, 85.
Füredi, A. A. and Valentine, R. C. (1962) Biochem. Biophys. Acta 56, 33.
Hartwell, L. H., Culotti, J., Pringle, J. R., and Reid, B. J.; (1974) Science 183, 46.
Hess, B., and Chance, B., (1978), Theo. Chem., 4, 159.
Holzapfel, C., Vienken, J. and Zimmermann, U. (1982) J. Membrane Biol. 67, 13.
Kaiser, F. (1982) Coll. Phenomena 5, in press;(this Volume, 1983).
McCann, F. V., Cole, J. J., Greyre, P. M., and Russell, J. A. G. (1983), Science 219, 991.
Mischel, M. and Lamprecht, I. (1980) Z. Naturforsch. 35c, 1111.
Mischel, M., Voss, A., and Pohl, H., (1982) J. Biol. Physics 10, 223.
Mischel, M. and Lamprecht, I. (1983) "Rotation of cells in nonuniform rotating alternating fields," J. Biol. Physics, 11, (in press).
Mischel, M. and Pohl, H. (1983) "Cellular spin resonance in rotating fields," J. Biol. Physics, 11. (in press).
Noyes, R. M. and Field, R. J. (1974) Ann. Rev. Phys. Chem. 25, 95.
Noyes, R. M. (1977) Acc. Chem. Res. 10, 214, and 273.
Okada, Y., Tsuchiya, W., and Inouye, A. (1979), J. Membrane Biol., 47, 357.
Pohl, H. (1951) J. Appl. Phys. 22, 869.
Pohl, H. and Crane, J. S. (1971) Biophys. J. 11, 711.
Pohl, H. and Wyhof, (1972) J. Non-Cryst. Solids, 11, 137.
Pohl, H. and Pollak, M. (1977) J. Chem. Phys. 66, 4031.
Pohl, H. (1978) "DIELECTROPHORESIS, The Behavior of Matter in Non-uniform Electric Fields", Cambridge University Press.
Pohl, H. (1980a) "Micro-dielectrophoresis of dividing cells," in BIOELECTROCHEMISTRY, pp. 273-95, edited by H. Keyzer and F. Gutmann, Plenum Press.
Pohl, H. (1980b), Int. J. Quantum Chem. 7, 411.
Pohl, H. A., (1980c) J. Biol. Phys. 7, 1.
Pohl, H. (1981a), J. Theor. Biol. 93, 207.
Pohl, H. (1981b), J. Bioenerg. Biomembranes 13, 149.
Pohl, H. Braden, T., Robinson, S., Piclardi, J., Pohl, D. G., (1981) J. Biol. Phys. 9, 133.
Pohl, H. (1982) Int. J. Quantum Chem. 9, 399.
Pohl, H. and Braden, T. (1982) J. Biol. Phys. 10, 17.
Pohl, H. (1983) "Cellular spinning in pulsed rotating electric fields," J. Biol. Physics, 11, (in press).
Rapp, P. E. (1979), J. Exp. Biol. 81, 281.
Rivera, H. (1983) private communication.
Rivera, H., Biscar, J. P., and Pohl, H. A., (1983) (private communication).

Rowlands, S., Sewchand, L. S., Lovlin, R. E., Beck, J. S., and Enns, E. G., (1981) Phys. Lett. 82A, 436.

Rowlands, S., Sewchand, L. S., Enns, E. G. (1982) Can. J. Physiol. Pharm. 60, 52.

Rowlands, S., (1982) J. Biol. Phys. 10, 199.

Rowlands, S., Eisenberg, C. P., and Sewchand, L. S. (1983), "Contractils: Quantum mechanical fibrils," J. Biol. Phys. 11, (in press).

Roy, S. C., Braden, T., and Pohl, H., (1981) Phys. Lett. 83A, 142.

Schmidt, S. and Ortoleva, P. (1979), J. Chem. Phys. 71, 1010.

Smith, Cyril W. (1983) (private communication).

Teixera-Pinto, A. A., Nejelski, L. L., Cutler, J. L., and Heller, J. H. (1960), Exp. Cell Res. 20, 548.

Treherne, J. E., Foster, W. A. and Schofield, P. K. (1979), "Cellular Oscillators" J. Exp. Biol. 81 (review volume).

Webb, S. J. (1980) Phys. Rep. 60, 201.

Westby, M., and Kirschbaum, F. (1982) J. Compar. Physiol., 145 , 399.

Yano, J., Okada, Y., Tsuchiya, W., Kinoshita, M., Tominaga, T., and Nishimura, M.; (1981), Acta Obst. Gyn. Japan, 33, 137.

Zimmermann, U., Vienken, J. and Pilwat, G. (1981), Z. Naturforsch. 36c, 173.

NONLINEAR DYNAMICS OF ORGANIC MOLECULES, INCLUDING BIOMOLECULES

THE ENERGETICS AND MOLECULAR DYNAMICS OF THE PROTON PUMPING PHOTOCYCLE IN BACTERIORHODOPSIN

Robert R. Birge,[a,b] Albert F. Lawrence,[b,c] Thomas M. Cooper,[b] Craig T. Martin,[d] David F. Blair[d] and Sunney I. Chan[d]

Department of Chemistry, University of California
Riverside, CA 92521 and A. A. Noyes Laboratory of
Chemical Physics, California Institute of Technology
Pasadena, CA 91125

ABSTRACT

The energy stored in the primary photochemical event of light adapted bacteriorhodopsin ($\sim$16 kcal mol^{-1}) is shown to be sufficient to pump two, but not three, protons per photocycle. A two channel model of the proton pumping photocycle is presented which accommodates the available experimental data. This model is energetically efficient because the primary event is used to destabilize both channels simultaneously. Two separate electrostatic gradients are generated, and the dark reactions that occur following the primary event restabilize both channels via proton translocation processes that ultimately pump two protons. The "primary" channel pumps the first proton in parallel with the formation of the M_{412} intermediate. The "secondary" channel pumps the second proton during the $O_{660} \rightarrow bR_{568}$ dark reaction provided the electrochemical gradient is below a given threshold. Ground state 13-cis→all-trans reisomerization is accomplished during the $N_{520} \rightarrow O_{660}$ dark reaction due to formation of a protonated Schiff base chromophore with a single (secondary) counterion near the β-ionylidene ring. This latter counterion stabilizes a chromophore resonance structure which has a high degree of bond

---

a) Address correspondence to this author

b) University of California, Riverside

c) Permanent address: Hughes Aircraft, P.O. Box 9399, Long Beach, CA 90810

d) California Institute of Technology

order reversal lowering significantly the barrier to ground state $C_{13}=C_{14}$ isomerization.

INTRODUCTION

Bacteriorhodopsin is the light harvesting protein of the purple membrane of Halobacterium halobium (for recent reviews see Refs. 1-5). The light adapted form of this protein undergoes a photochemical cycle that is responsible for transporting protons from the inside (cytoplasm) to the outside (extracellular) of the membrane (Fig. 1). The resulting hydrogen ion gradient is used to chemiosmotically drive the synthesis of ATP. Halobacterium halobium grows the purple membrane under conditions of oxygen deprivation when the normal respiration process becomes too inefficient to sustain growth.

The purple membrane pumps ∿2 protons per photocycle under the pH and ionic strength conditions inherent in the ambient environment of this halophilic microorganism.[1,6-7] This observation indicates that the simple notion that the retinyl Schiff base proton is translocated as a result of electrostatic forces generated by chromophore isomerization is, by itself, insufficient to explain the pumping mechanism. A more complex model which accommodates a larger stoichiometry (2 protons pumped/photocycle) and the photocycle energetics[8] is required. The purpose of this paper is to present a plausible, two-channel model of the proton pumping cycle which accommodates the available experimental evidence. We note at the outset of this discussion that the proposed model is not unique in that other mechanisms are possible. Accordingly, we present this model as a preliminary mechanistic study with the hope that it will stimulate further experimental and theoretical study of this complex biological process.

ENERGETICS OF THE PROTON PUMPING CYCLE

The photochemical cycle of light adapted bacteriorhodopsin is shown in Fig. 1. The primary event ($bR_{568} \rightarrow K_{610}$) is the only molecular process which occurs photochemically. The complex cycle that follows the formation of K_{610} is thermal and is therefore constrained to utilize the energy stored in K_{610}. A recent laser photocalorimetric measurement of energy storage in bacteriorhodopsin indicates that 15.8 ± 2.5 kcal mol^{-1} of energy is stored in the primary photoproduct.[8] Steady state illumination of intact cells induces the protein to eject protons from the cytoplasm generating an electrochemical gradient across the cell membrane. This gradient is partitioned into two components,[9]

$$\Delta p = \Delta\psi - 2.3RT\Delta pH/F \qquad (1)$$

$$= \Delta\psi - 59\Delta pH \text{ (mV)} \qquad (25°C) \qquad (2)$$

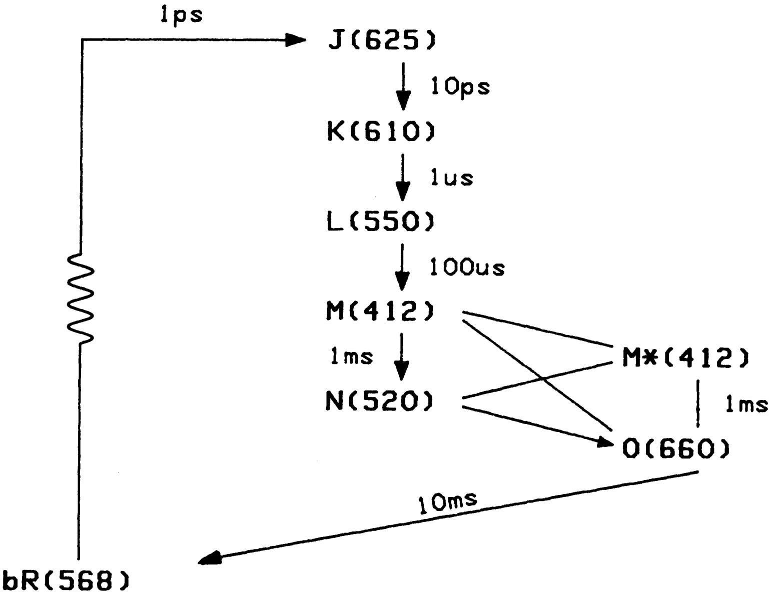


Fig. 1. The photochemical cycle of light-adapted bacteriorhodopsin. Relative free energies are approximately related to vertical position in the figure. The formation times of the intermediates are order of magnitude estimates for ambient temperature. The approximate absorption maxima (nanometers) of the intermediates are indicated in parentheses and a given intermediate will be designated in the text using these values as subscripts [i.e. K_{610} represents K(610)]. There is no consistency in the literature concerning the absorption maxima assignments for the intermediates. Accordingly, the primary photoproduct is variously labelled as K_{590}, K_{610} and K_{630} in the literature. Although a short-lived transient species, J_{625}, has been observed prior to K_{610}, we will refer to the primary event as the $bR_{568} \rightarrow K_{610}$ reaction and ignore the J_{625} intermediate. The exact chronology of the dark reactions is not precisely known, and we will assume the reaction sequence bR→K→L→M→N→O→bR is the primary proton pumping sequence in light-adapted bacteriorhodopsin (see text).

where Δp is the proton motive force (or electrochemical gradient), $\Delta\psi$ is the electrical potential difference and ΔpH is the pH gradient across the membrane. Equation 2 is obtained by evaluating RT/F (R = gas constant, F = Faraday constant) at T = 298K. The pH gradients during illumination are roughly in the range 0.7 to 1.5 units (inside alkaline), but the value is subject to considerable uncertainty and is dependent upon extracellular pH.[10] Estimates of $\Delta\psi$ during illumination generally fall in the range -120 to -220 mV.[10] Accordingly, the electrochemical gradient across the cell may reach ∿300 mV, but for the purposes of this discussion we will assume a gradient of 250 mV (= $-\Delta p$). The energy required to pump two protons across a gradient of 250 mV is ∿12 kcal mol^{-1} (1 kcal mol^{-1} = 43.4 mV/molecule) which implies that the energy stored in the primary event (∿16 kcal mol^{-1}) is sufficient to pump two but not three protons per photocycle. The observed stoichiometry of 2 protons/photocycle implies a very high efficiency for the proton pumping mechanism of bacteriorhodopsin of ∿75% (only ∿4 kcal mol^{-1} of internal energy is lost per photocycle). Although this efficiency is subject to considerable error (a lower estimate of Δp of -150 mV is within experimental error), the above analysis suggests that a majority of the energy stored in the primary event is utilized to generate the electrochemical gradient. One of the primary objectives of our model is to provide a possible explanation of this high efficiency.

EXPERIMENTAL BACKGROUND

Our model of the proton pumping cycle of bR_{568} is based on the energetic requirements described in the previous section as well as a number of experimental observations. The available literature on this subject is extensive and the following observations briefly summarize the salient experimental data:

(1) There are two counterions associated with the retinyl protonated Schiff base (RPSB) chromophore in bR_{568}. The primary counterion is located ∿3Å from the Schiff base nitrogen.[11-14] A second counterion is located above the plane of the β-ionylidene ring ∿3.5Å from C_5.[11-12] Our model of the retinyl chromophore, its protonated Schiff base linkage with lysine 216 and the primary counterion (Aspartate 212) is shown in Fig. 2.[15] A tyrosine residue (Tyr 26) is also located near the RPSB chromophore.[16,17]

(2) A tyrosine residue is deprotonated[18-19] and a carboxylate group is protonated[20-22] during the photocycle. Both events appear to occur on a time scale identical to the formation time of the M_{412} intermediate within experimental error. A second carboxylate group may also be protonated at a slightly slower rate (subsequent to M_{412} formation) during the photocycle, and both of these groups are subsequently deprotonated during the $N_{520} \rightarrow O_{660} \rightarrow bR_{568}$ dark reactions.[20]

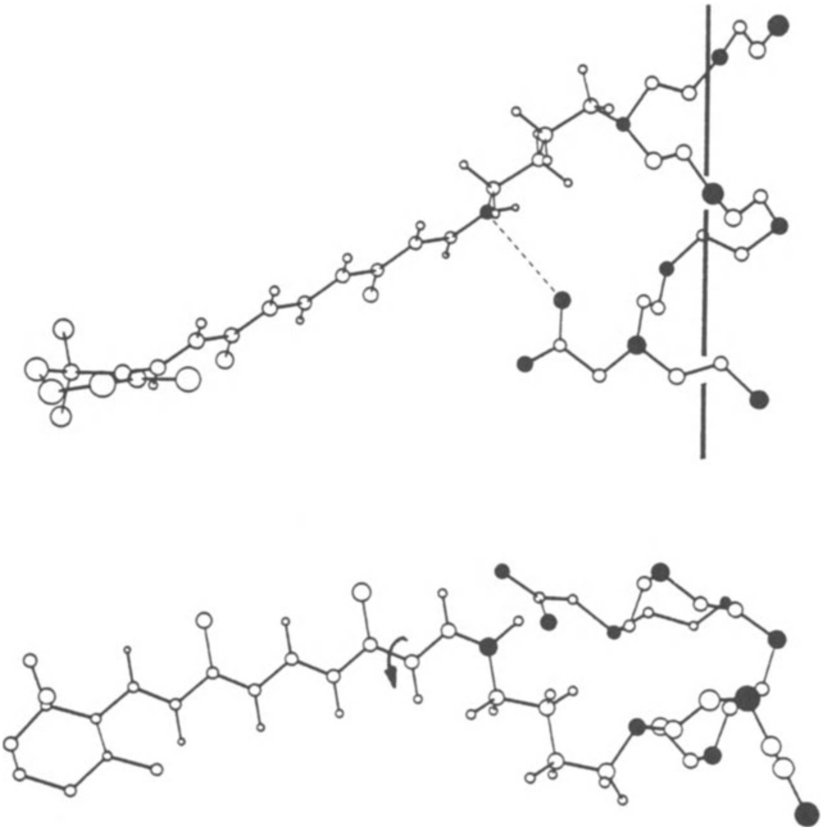


Fig. 2. Our model of the chromophore binding site in light adapted bacteriorhodopsin (bR$_{568}$). The all-*trans* retinyl chromophore and the local portion of the α-helix to which it is bound is shown from two perspectives. In the top picture, the site is viewed from inside the protein so that the trans-membrane axis of the helix is in the plane of the paper (the thick solid line indicates the helix axis). In the bottom picture, the site is viewed from the top of the protein from the inside (cytoplasmic) membrane surface. The chromophore is bound to the protein backbone via a protonated Schiff base linkage with lysine (Lys 216). The primary counterion is Aspartate 212, which is separated from the Schiff base nitrogen by approximately 3Å (dotted line in top picture). The other amino acids attached to the protein backbone are not shown. The methyl groups attached to the chromophore are indicated using open circles slightly larger than those used to indicate the hydrogen atoms. The primary photochemical event involves a 13-*trans* to 13-*cis* photoisomerization of the chromophore away from the counterion [out of the plane of the paper in the bottom figure (see arrow)].

(3) The primary photochemical event ($bR_{568} \rightarrow K_{610}$) involves an all-trans to (distorted) 13-cis photoisomerization (see Ref. 2 and references therein). Chemical extraction techniques have demonstrated that L_{550} and M_{412} contain 13-cis chromophores.[23,24] Recent resonance Raman experiments indicate the O_{660} has a distorted all-trans geometry.[25] Accordingly, 13-cis to all-trans isomerization occurs during the $M_{412} \rightarrow N_{520} \rightarrow O_{660}$ dark reactions. It should be noted that the N_{520} intermediate is "poorly resolved"[1] and many investigators ignore this intermediate entirely. Furthermore, rapid equilibration among the latter intermediates (see Fig. 1) is postulated by many investigators (see discussion in Ref. 1). The chromophore is protonated in all of the intermediates with the exception of M_{412}.[2]

A TWO CHANNEL MODEL OF THE PROTON PUMPING PHOTOCYCLE

Our model of the proton pumping photocycle of light adapted bacteriorhodopsin is shown in Fig. 3. This model was designed to accommodate the available experimental evidence (see above), but it is not a unique model in that other reaction schemes are possible. Several treatments of the theory of proton transport in hydrogen bonded chains have recently been published.[26] A principal feature which distinguishes this model from others which might be proposed is the involvement of two proton conducting channels. The presence of two channels allows two protons to be translocated in parallel rather than in series. The high efficiency of the proton pumping mechanism must be mechanistically accommodated, and a single channel model is less efficient for two reasons. The most significant problem with a single channel model is that initiation of the second proton translocation requires a reciprocating motion which is inherently inefficient. A two channel model has the further advantage that a single light activated step can destabilize both channels simultaneously. Hence, energy storage inherent in the primary event can be used efficiently to generate two gradients. The dark reactions that occur subsequently restabilize these channels via proton translocation processes that ultimately pump two protons.

The model shown in Fig. 3 is intentionally ambiguous concerning the specific assignments of a majority of the residues associated with the two channels. Furthermore, the use of seven residues in both channels is arbitrary and probably underestimates the number of residues in each channel by a factor of two. We note also that the channels may in fact converge or share common residues in regions separated from the chromophore binding site. This possibility would not detract from the aforementioned advantages of a two channel model. The two channels come into fairly close contact at the chromophore which is located near the center of the membrane.[1,3] The two channels are in "electrostatic contact" in the vicinity of the chromophore, which is required for stabilization of some of the intermediates (see below).

We represent the entire chromophore (Fig. 2) using the symbol "C-N-H" to represent the all-*trans* conformation and the symbol "H-N-C" to represent the 13-*cis* conformation. The unprotonated 13-*cis* chromophore in M_{412} is designated "N-C". The primary counterion (Aspartate 212)(Fig. 2) is indicated by "O-A" (Fig. 3) and is part of the RHS channel. The secondary counterion near the β-ionylidene ring of the chromophore[11,12] is not shown, but the interaction of this counterion with the chromophore is important during the latter stages of the cycle (see below)

The left channel is designated as the "primary channel" because the proton translocated via this channel is pumped first and the extracellular ejection of this proton coincides with the formation of M_{412} and deprotonation of the chromophore. The right channel is the "secondary channel" because the proton translocated via this channel is pumped second (during the $O_{660} \rightarrow bR_{568}$ dark reaction) and is only pumped under favorable extracellular conditions. Accordingly, our model provides for the observed variable stoichiometry of the proton pumping cycle (see below).

The Primary Event ($bR_{568} \rightarrow K_{610}$)

The primary event involves an extremely rapid[27] all-*trans* to 13-*cis* photoisomerization of the chromophore.[2,15,23] The observed deuterium isotope effect suggests that a proton translocation may also occur, but whether or not this translocation involves a residue in either channel is, of course, unknown. The spectroscopic and energetic evidence suggests that this rapid proton translocation during the primary event does not involve the chromophore proton nor the primary counterion.[15] Although it would certainly be possible to incorporate a proton translocation associated with the primary

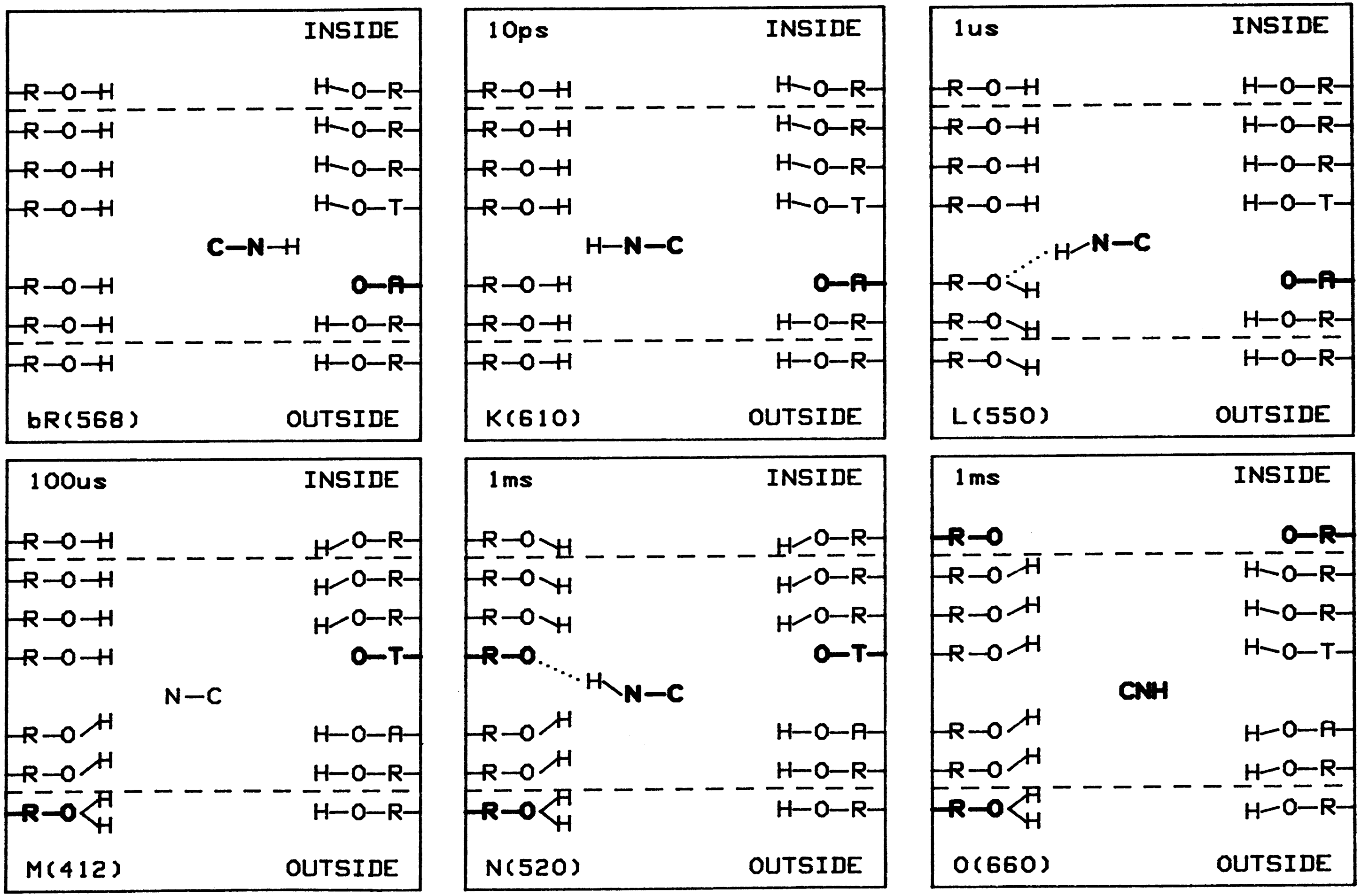

INSIDE
C—N—H
bR(568)
OUTSIDE
10ps
INSIDE
H—N—C
K(610)
OUTSIDE
1us
INSIDE
H—N—C
L(550)
OUTSIDE
100us
INSIDE
N—C
M(412)
OUTSIDE
1ms
INSIDE
H—N—C
N(520)
OUTSIDE
1ms
INSIDE
CNH
O(660)
OUTSIDE
R—O—H
H—O—R
H—O—T
O—R
O—T

Fig. 3. A schematic diagram of the proposed two channel model of the proton pumping mechanism in light-adapted bacteriorhodopsin. The various intermediates are assigned in the lower left-hand corner of the rectangular boxes, and order of magnitude estimates of their formation times following excitation are indicated in the upper left-hand corner. The primary counterion (ASP 212) and Tyrosine 26 are indicated by "O-A-" and "O-T-", respectively. All the other proton donors and acceptors are indicated by "-R-O" (left channel) or "O-R-" (right channel). The protonated Schiff base chromophore is indicated by "C-N-H" (all-*trans*) or "H-N-C" (13-*cis*). Residues which carry a net charge are indicated in bold-face.

event into our model, we doubt that it is a significant component of the proton pumping cycle and we therefore do not attempt to assign this event to any specific process.

The photoisomerization is the primary source of the energy storage in the primary event.[8,15] The experimental[8] and theoretical[15] evidence suggests that approximately half of the energy storage is associated with charge separation, and half is associated with conformational distortion of the chromophore-protein ensemble. The movement of charge from the right channel to the left channel simultaneously destabilizes both channels and it is the restabilization of these channels during the dark reactions that is responsible for pumping two protons.

The $K_{610} \rightarrow L_{550}$ Dark Reaction

This reaction occurs in ∿1 μs and is viewed as a partial relaxation of the left channel involving small motions of the residues near the chromophore and the generation of a weak hydrogen bond involving the chromophore proton and a local carboxylate or other hydroxy residue. The formation of the hydrogen bond stabilizes the ground state and blue shifts the absorption maximum. A portion of this blue shift is due to chromophore relaxation.

The $L_{550} \rightarrow M_{412}$ Dark Reaction

The hydrogen bond formed during the preceeding $K_{610} \rightarrow L_{550}$ reaction is viewed as the precursor to the present reaction which stabilizes the primary channel via deprotonation of the chromophore to form M_{412}. This reaction stabilizes the entire primary channel by returning the channel to its original hydrophobic state and results in the ejection of the first proton into the extracellular medium. A simultaneous reaction in the secondary (RHS) channel occurs to stabilize the primary counterion (Asp 212) which was originally destabilized during the primary event. The proton donor is a tyrosine (Tyr 26?) which is preferentially deprotonated due to local pK_a changes. In this regard, our model follows previous models based on light-induced pK_a changes in the protein.[28,29] Merz and Zundel have suggested that Tyrosine 26 and Aspartate 212 are in close proximity,[30] and the transfer of a proton from a tyrosine to a carboxylate group during M_{412} formation is experimentally observed.[18-22]

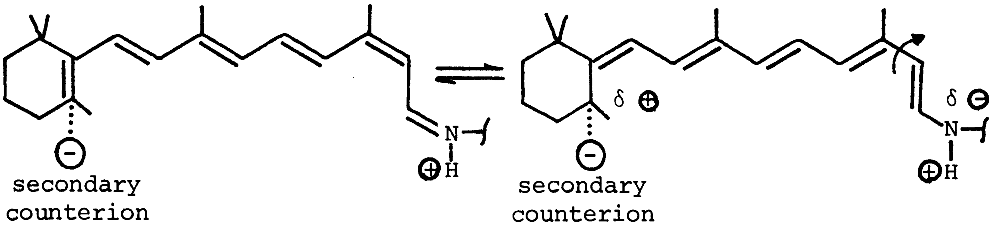
The stabilization of the left and right channels has destabilized the chromophore. The secondary counterion is presumed to remain unmasked which places the β-ionylidene ring in an hydrophilic environment which is energetically unfavorable for the Schiff base. Subsequent dark reactions now take place to stabilize the chromophore via hydrogen bonding and eventual reprotonation.

The $M_{412} \rightarrow N_{520}$ Dark Reaction

This dark reaction is relatively slow and involves "partial" reprotonation of the retinyl chromophore to form a strongly hydrogen bonded species with the hydrogen shared between the chromophore and a primary channel residue. The N_{520} intermediate is very similar to the L_{550} intermediate except that the hydrogen bonding residue is a different group in the primary channel (Fig. 3). Note that in both L_{550} and N_{520}, the primary channel hydrogen bonding residue is across from the secondary channel counterion (Asp 212 in L_{550}, Tyr 26 in N_{520}). This provides electrostatic stabilization of the hydrogen bonded species and assumes a certain degree of flexibility in the movement of the chromophore. It is perhaps more reasonable to view the $M_{412} \rightarrow N_{520}$ dark reaction as an equilibrium ($M_{412} \rightleftarrows N_{520}$) because the driving force for this step, and the barrier for this interconversion, are probably very small.

The $N_{520} \rightarrow O_{660}$ Dark Reaction

One of the interesting, unsolved problems associated with the proton pumping cycle of bacteriorhodopsin is how the chromophore reisomerizes in the dark to form the all-trans conformation. We believe the key to this mechanism is the $N_{520} \rightarrow O_{660}$ dark reaction. Resonance Raman experiments indicate that the chromophore in O_{660} is distorted all-trans. Theoretical calculations indicate that the barrier to 13-cis→13-trans double bond isomerization in the ground state is much lower in the protonated Schiff base than in the unprotonated Schiff base.[31] This barrier decreases further for a protonated Schiff base in the absence of the primary counterion due to the importance of the resonance structure shown on the right-hand-side of the following diagram:[31]



The tyrosine residue in N_{520} provides a distant primary counterion that favors the left-hand resonance structure. However, the translocation of protons down the secondary channel to mask the tyrosine counterion immediately favors this right-hand resonance structure and the chromophore immediately isomerizes back to the all-trans conformation. The 13-cis to all-trans isomerization relieves ∿1.3 kcal mol^{-1} of strain energy in the chromophore.[32] The extremely red shifted absorption maximum is characteristic of a retinyl protonated Schiff base in the absence of a primary counterion.[2]

The O_{660} to bR_{568} Dark Reaction

The second proton is pumped during this slow (∼10 ms) dark reaction to reform bR_{568}. The driving force for this reaction is the electrostatic stabilization of the protonated Schiff base via the unmasking of the primary counterion (Asp 212). This event pumps the final proton via the secondary channel provided the electrochemical gradient is not too large. The alternative (which will not result in an extracellular proton being pumped) is deprotonation of the primary counterion through transfer of a proton back "up" the secondary channel (see Fig. 3). The isomerization of the chromophore will selectively translocate the proton in the direction of the double bond rotation. In other words, a clockwise 13-cis→13-trans isomerization in O(660)(Fig. 3) will select extracellular proton ejection. Our model presumes isomerization is constrained by the protein to preferentially choose this pathway. However, a large electrochemical gradient could override this preference and we view the O_{660}→bR_{568} dark reaction as a weak proton pumping reaction step.

COMMENTS AND CONCLUSIONS

The proton pumping mechanism of light-adapted bacteriorhodopsin is extremely complex and the present model is surely inadequate to fully explain the intricacies of this process. Two elements of our model, however, deserve closer scrutiny because they have important advantages which have been lacking in previous models. These elements are, 1) the use of two separate proton conducting channels and 2) the concept that reisomerization occurs due to the interaction of the chromophore with its secondary counterion during masking of its primary counterion. The two channel model is energetically efficient because the primary event destabilizes both channels simultaneously. Two separate electrostatic gradients are therefore developed simultaneously, and the dark reactions that occur following the primary photochemical event restabilize both channels via dual extracellular proton ejection. A plausible explanation for the ground state reisomerization of the chromophore (13-cis→all-trans) is advanced based on the assumption that this process occurs when the protonated Schiff base chromophore is stabilized solely by a secondary counterion near the β-ionylidene ring C_5 carbon atom. Resonance stabilization of this structure results in significant ground state bond order reversal which generates a pseudo C_{13}-C_{14} single bond thereby lowering the barrier to reisomerization. The proposed structure provides a realistic explanation of the highly red-shifted absorption maximum of the O_{660} intermediate.

REFERENCES

1. W. Stoeckenius and R.A. Bogomolni, Annu. Rev. Biochem. 51:587 (1982).
2. R.R. Birge, Annu. Rev. Biophys. Bioeng. 10:315 (1981).
3. L. Packer, ed., "Visual Pigments and Purple Membranes," in Methods in Enzymology, vol. 88, Academic, New York (1982).
4. R.R. Alfano, ed. "Biological Events Probed by Ultrafast Laser Spectroscopy," Academic, New York (1982).
5. M. Ottolenghi, Adv. Photochem. 12:97 (1980).
6. R. Govindjee, T.G. Ebrey and A.R. Crofts, Biophys. J. 30:231 (1980).
7. M. Renard and M. Delmelle, Biophys. J. 32:993 (1980).
8. R.R. Birge and T.M. Cooper, Biophys. J. 42:61 (1983).
9. S.J. Ferguson and M.C. Sorgato, Annu. Rev. Biochem. 51:185 (1982).
10. J.K. Lanyi, Microbiol. Rev. 42:682 (1978).
11. K. Nakanishi, V. Balogh-Nair, M. Arnaboldi, K. Tsujimoto and B. Honig, J. Am. Chem. Soc. 102:7945 (1980).
12. V. Balogh-Nair, J.D. Carriker, B. Honig, B. Kamat, M.G. Motto, K. Nakanishi, R. Sen, M. Sheves, M.A. Tanis and K. Tsujimoto, Photochem. Photobiol. 33:483 (1981).
13. U.C. Fischer and D. Oesterhelt, Biophys. J. 31:139 (1980).
14. A. Warshel and M. Ottolenghi, Photochem. Photobiol. 30:291 (1979).
15. R.R. Birge and B.M. Pierce, Laser Chem. (in press).
16. H.D. Lemke and D. Oesterhelt, Eur. J. Biochem. 115:595 (1981).
17. V. Rosenbach, R. Goldberg, C. Gilon and M. Ottolenghi, Photochem. Photobiol. 36:197 (1982).
18. O. Kalisky, M. Ottolenghi, B. Honig and R. Korenstein, Biochemistry 20:649 (1981).
19. B. Hess and D. Kuschmitz, FEBS Letters 100:334 (1979).
20. F. Siebert, W. Mantele and W. Kreutz, FEBS Letters 141:82 (1982).
21. K. Bagley, G. Dollinger, L. Eisenstein, A.K. Singh and L. Zimanyi, Proc. Natl. Acad. Sci. USA 79:4972 (1982).
22. K.J. Rothschild, M. Zagaeski and W.A. Cantore, Biochem. Biophys. Res. Commun. 103:483 (1981).
23. M. Tsuda, M. Glaccum, B. Nelson and T.G. Ebrey, Nature (London) 287:351 (1980).
24. M.J. Pettei, A.P. Yudd, K. Nakanishi, R. Henselman and W. Stoeckenius, Biochemistry 16:1955 (1977).
25. S.O. Smith, J.A. Pardoen, P.P.J. Mulder, B. Curry, J. Lugtenberg and R. Mathies, Biochemistry (in press).
26. J.F. Nagle, M. Mille and H. J. Morowitz, J. Chem. Phys. 72:3959 (1980); Y. Kashimori, T. Kikuchi and K. Nishimoto, J. Chem. Phys. 77:1904 (1982).
27. M.L. Applebury, K.S. Peters and P.M. Rentzepis, Biophys. J. 23:375 (1978).
28. O. Kalisky, M. Ottolenghi, B. Honig and R. Korenstein, Biochemistry 20:649 (1981).

29. S. Druckmann, M. Ottolenghi, A. Pande, J. Pande and R.H. Callender, Biochemistry 21:4953 (1982).
30. H. Merz and G. Zundel, Chem. Phys. Letters 95:529 (1983).
31. K. Schulten and P. Tavan, Nature 272:85 (1978).
32. R.R. Birge and B.M. Pierce (to be published).

PRESENCE OF NONLINEAR EXCITATIONS IN DNA STRUCTURE AND THEIR RELATIONSHIP TO DNA PREMELTING AND TO DRUG INTERCALATION

Asok Banerjee and Henry M. Sobell

Department of Radiation Biology and Biophysics
The University of Rochester
School of Medicine and Dentistry
Rochester, New York 14624

Department of Chemistry
The University of Rochester River Campus Station
Rochester, New York 14624

ABSTRACT

We propose that collectively localized nonlinear excitations (solitons) exist in DNA structure. These arise as a consequence of an intrinsic nonlinear ribose inversion instability that results in a modulated β alternation in sugar puckering along the polymer backbone. In their bound state, soliton- antisoliton pairs contain premelted core regions capable of undergoing breathing motions that facilitate drug intercalation. We call such bound state structures β premeltons. The stability of a β premelton is expected to reflect the collective properties of extended DNA regions and to be sensitive to temperature, pH, ionic strength and other thermodynamic factors. Its tendency to localize at specific nucleotide base sequences may serve to initiate site-specific DNA premelting and melting. We suggests that β premeltons provide nucleation centres important for RNA polymerase-promoter recognition. Such nucleation centers could also correspond to nuclease hypersensitive sites.

The possibility that nonlinear excitations (solitons) exist in biopolymers and play a central role in energy transfer was first advanced by Davydov in his classic series of papers (1-3). In addition, a different class of solitons that give rise to localized conformational changes in DNA structure has been proposed by Englander et al. to explain DNA breathing phenomena (4).

Solitons are intrinsic locally coherent excitations that move along a polymer chain with a velocity significantly less than the speed of sound (they may even be stationary). They are combinations of intramolecular and deformational excitations that appear as a consequence of an intrinsic nonlinear instability in the polymer structure.

Extensive research on solitons in many physical systems has shown that this nonlinearity gives the spatially localized conformational excitation a robust character (5,6). Solitons do not significantly interact with conventional normal mode excitations (i.e., phonons). They have their own identity and can be treated by Newtonian dynamics as heavy Brownian-like particles, each having an "effective mass". Solitary structures -- as sites for biochemical activity -- behave like independent species and can be treated by statistical mechanics and chemical thermodynamics. They can arise from equilibrium or nonequilibrium processes.

Here, we propose that localized nonlinear excitations -- solitons -- exist in DNA. These arise as a consequence of an intrinsic nonlinear instability in DNA structure which is primarily associated with interconversions between the two predominant sugar-pucker conformations, C2' _endo_ and C3' _endo_. In their bound state, soliton-antisoliton pairs surround β premelted core regions -- these regions can undergo breathing motions that facilitate the intercalation of drugs and dyes into DNA. Similar bound state structures could act as phase coundaries that connect different DNA forms.

Several communications describing our ideas have already appeared (7-9). Here, we develop these ideas in greater detail.

Drug Intercalation into DNA

Structural features of our theory have been suggested by information obtained through X-ray crystallographic studies of model drug-DNA complexes. These studies have indicated that DNA undergoes a highly specific conformational change at the immediate intercalation site when binding a class of intercalating agents known as simple intercalators. This (localized) structural state is characterized in part by the presence of a mixed sugar-puckering pattern connecting adjacent base-paired nucleotides -- C3' _endo_ (3'-5') C2' _endo_ -- and, in addition, involves less pronounced changes in the torsional angles that describe the glycosidic and sugar-phosphate bond linkages. This allows DNA to stretch and to unwind when accommodating an intercalator.

More generally, we have proposed this (base-paired dinucleotide) structure to belong to a family of structures, each of which corresponds to an element of organized β structure (see discussion below) (10). All elements possess the same mixed-sugar puckering pattern and have similar backbone conformational angles -- but vary in the degree of base-unstacking. Lower energy forms contain base-pairs completely unstacked. Steroidal diamines such as irehdiamine A, bind to the lower energy β element through partial intercalation, while planar drugs and dyes bind to higher energy β elements by complete intercalation (11).

A DNA structure of particular interest is stabilized at saturating concentrations of drugs such as irehdiamine and ethidium. This is an organized helical structure that contains the β element as the asymmetric unit. We call this neighbor-exclusion structure β DNA, or, for reasons to be described shortly, β premelted DNA (12).

Irehdiamine stabilizes a low energy form of β premelted DNA, while ethidium stabilizes a higher energy form. We have described the molecular nature of these structures elsewhere.

Kink-Antikink Bound States in DNA Structure -- the β Premelton

How do β structural elements arise in DNA -- and what are the surrounding structural features on either side of these elements?

We propose that β elements arise as part of more organized β structure, whose appearance reflects the presence of a soliton-antisoliton (i.e., kink- antikink) bound state in DNA structure (see Figures 1a and 1b). Such a structure contains a modulated β alternation in sugar puckering about the central β premelted core region which gradually merges into B DNA on either side -- typically, a structure of this kind would give rise to an energy density profile similar to that shown in Figure 1c (13). We call this type of composite structure a β premelton (14).

We have already described the methods used to construct this β premelton (9). The B DNA to β DNA transition was first computed as a homogeneous transition involving the entire polymer length. This was accomplished in a series of steps in which the sugar puckering of alternate deoxyribose residues was altered and the structure then energy minimized subject to a series of constraints and restraints. To simulate the bound state structure, base-paired dinucleotide elements from each structure in this sequence

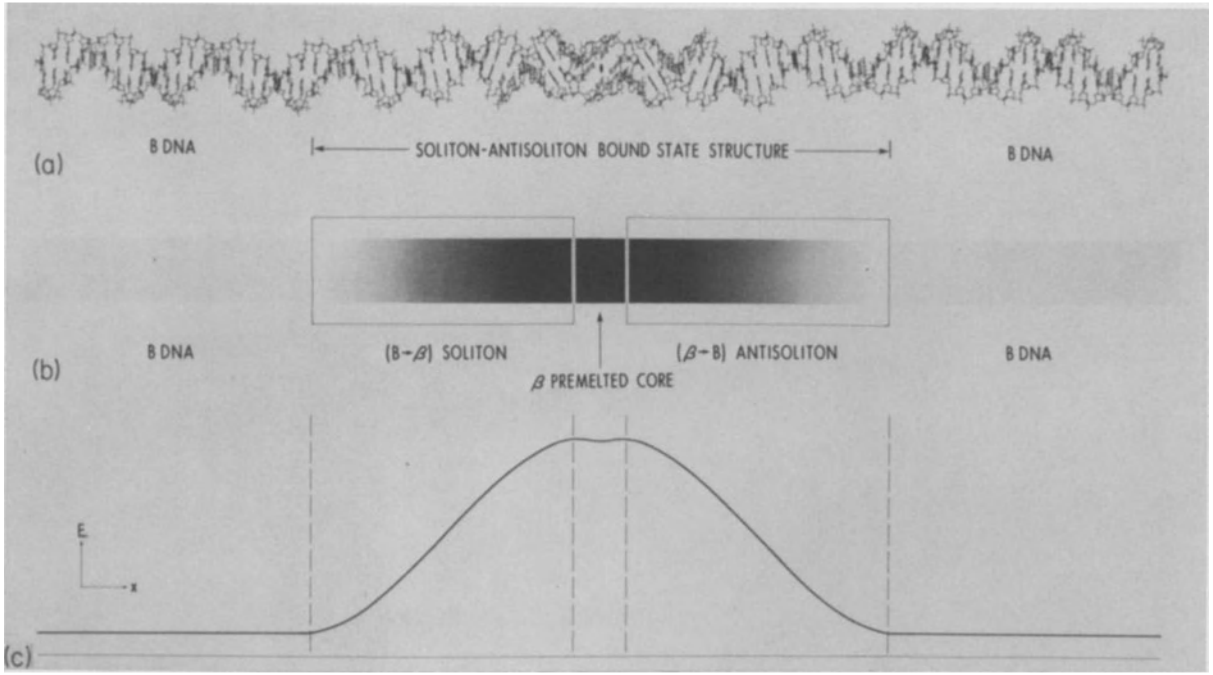


Figure 1. (a) Molecular representation of the soliton-antisoliton bound state structure in B DNA. (b) Schematic illustration of the gradation of structural change that connects the β premelted core with B DNA on either side through soliton-antisoliton pairs. (c) A typical energy density profile of the kink-antikink bound state structure -- the dip in the center of the energy density profile signifies the presence of a metastable structural state. We call this type of composite structure a β premelton.

were then pieced together using a least squares procedure. The decision to compute twelve intermediate structures and to then use these to construct the soliton- antisoliton bound state (without a detailed knowledge of the total energies involved) is somewhat arbitrary, but does not alter the basic conclusions presented here. More complete computations are in progress to obtain a family of minimum energy bound state structural forms.

We cannot say at the present time whether the center of the β premelton is open enough to accommodate an intercalator such as ethidium. However, structures such as these have an intrinsic ability to undergo low frequency breather motions and it is possible that such motions facilitate the intercalation process.

Mechanism of Drug Intercalation and Its Relationship to DNA Premelting

We envision intercalation to begin with the loose (external) binding by an intercalator to the B helix. A β premeltin then appears in the region -- this either immediately permits the drug molecule to intercalate, or a breathing motion of sufficient amplitude is first required. Simple intercalators may pin (immobilize) the β premelton. They could also prevent soliton-antisoliton anihilation (a spontaneous event that returns DNA in the immediate region to its ground state structute, i.e., B DNA). This is because simple intercalators bind the β premelted core region tightly. Complex intercalators may vary in this regard. Actinomycin could stabilize the β premelted core within the β premelton, pinning the structure and also preventing its self-anihilation (15). Daunomycin, however, may be unable to do this and ends up, therefore, binding to a conformation very similar to B DNA (16).

The term "soliton" is synonomous with "kink", and refers to a gradation of structural change either within a single structural type (i.e., a nontopological soliton) or connecting two different structures (i.e., a topological soliton). The structure shown in Figure 1a consists of two topological solitons -- a (B→β) soliton and a (β → B) antisoliton. These two structures surround the premelted core region and are related by 2- fold symmetry.

In Figure 1a, we show a β premelton containing a small premelted core. Premeltons having higher energies can occur. Bifunctional intercalators, such as echinomycin, require β premeltons with larger cores. Intercalators that necessitate the transient rupture of hydrogen bonds connecting base- pairs to gain entrance into (and out of) DNA (i.e. meso-tetra-(4-N- methylpyridyl) porphine) (17) require β premeltons with still larger cores. We term distortions that combine base unstacking with the transient rupture of hydrogen bonds connecting base-pairs -- DNA breathing distortions. We expect such (dynamic) distortions to take place in the center of extended regions of β premelted DNA.

The β premelted core region within soliton-antisoliton bound state structures can be thought of as a transition state that nucleates the DNA melting process (hence, the name β premelted DNA). At lower temperatures, soliton-antisoliton pairs surround small β premelted cores. As the temperature rises, these bounding soliton-antisoliton pairs move apart leaving central regions of the growing β premelted cores more and more disordered. Finally, with increasing temperature, denaturation bubbles appear that contain regions of denatured DNA connected to native DNA. These kink and antikink structures, therefore, act as structural boundaries connecting denatured DNA with native DNA in the helix to coil phase transition. They also act as (partial) phase boundaries in a variety of DNA structural phase transitions -- we will describe the nature of one such transition below.

B-A Junction

Soliton concepts allow us to propose the molecular nature of the junction that connects B with A DNA during the B to A structural phase transition. Such a junction (shown in Figures 2a and 2b) has been constructed in a manner similar to that described above.

A homogeneous β DNA to A DNA change was computed as a sequence of transitions involving the entire polymer length. Dinucleotide elements from this sequence and the B to β sequence were then connected together using the least squares joining procedure. The final structure consists of two different topological solitons -- a (B → β) soliton and a (β → A) antisoliton and vice versa. Such a junction constitutes a phase boundary transforming B DNA to A DNA. In the global sense, this kink-antikink bound state is topological since it alters the structure of the polymer as it moves along its length.

How does the B to A transition begin?

We envision the B to A transition to begin at the centers of premeltons with B DNA structure. Nucleation could be site specific, and involve growing regions of A DNA joined to B DNA on either side as [(B →β): (β → A)] and [(A → β): (β → B)] β premelton pairs (see Figures 3a - 3d). Each β premelton contains a small β premelted core, and moves along the polymer with minimal activation energy. Its high free energy and ease of movement could account for the cooperative nature of the phase transition.

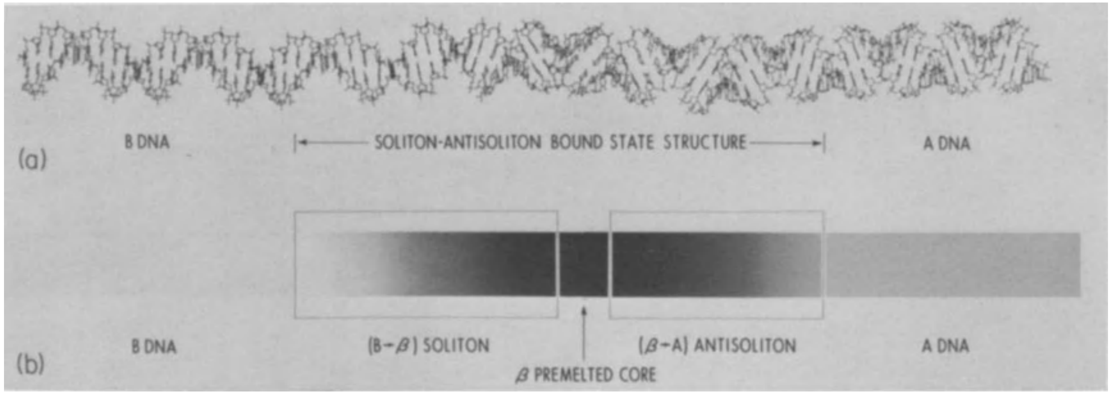


Figure 2. (a) Phase boundary connecting B with A DNA. (b) Schematic illustration of the gradation of structural change that connects the β premelted core region with B DNA on one side and A DNA on the other side through kink-antikink pairs. See text for discussion.

Discussion

Soliton concepts provide a strong rationale for expecting coherent nonlinear excitations to extend over multiple base pairs to provide either transient or permanent conformational changes in specific DNA regions. These will always be present to a certain extent at normal thermal energies, but their concentrations will depend on temperature, pH, ionic strength, hydration, extent of super helicity and other thermodynamic factors.

The stability of soliton-antisoliton pairs in naturally occurring DNA is further expected to reflect the detailed nucleotide base sequence. This is partly because the ease with which β premelted core regions form primarily reflects the magnitude of localized base stacking energies. Nucleotide sequences with minimal base overlap (i.e., alternating purine-pyrimidine sequences) are likely to be favored to form β premelted core regions, along with sequences that contain high A-T/G-C base ratios (18). Equally important are the energetics in the kink and antikink regions. As seen in Figure 1c, the stability of a kink-antikink bound state reflects the depth in the energy minimum within the β premelted core coupled with the height and separation of the domain walls on either side.

It is evident that sequentially homogeneous DNA polymers are not good models to understand the properties of naturally occurring DNA since they lack sufficient information in their nucleotide base sequence to give rise to this site specificity. In the context of soliton models, this poses the problem of describing soliton behavior in the presence of locally altered potentials -- a general theory for this has been developed (19). This theory shows that (nontopological) solitons either move nonuniformly or are trapped by locally favorable potentials. It remains to extend this theory to DNA structure to predict the localization of kink-antikink bound states at specific nucleotide sequences. Since the kink-antikink pair is multiple base-pairs in extent, it may be that the effective trapping potential involves the recognition of an extended sequence rather than being determined by any single base-pair energetics or its immediate neighbors.

We note experimental evidence that indicates the presence of nuclease hypersensitive sites in eukaryotic DNA, many of these located at 5' ends of genes (20,21). These same sites are sensitive to cleavage by a 1,10- phenanthroline-copper(I) complex, a known intercalating agent. It is possible that these sites correspond to the centers of β premeltons localized in these regions.

The presence of β premelted core regions within kink-antikink pairs could provide a key component in the recognition of the

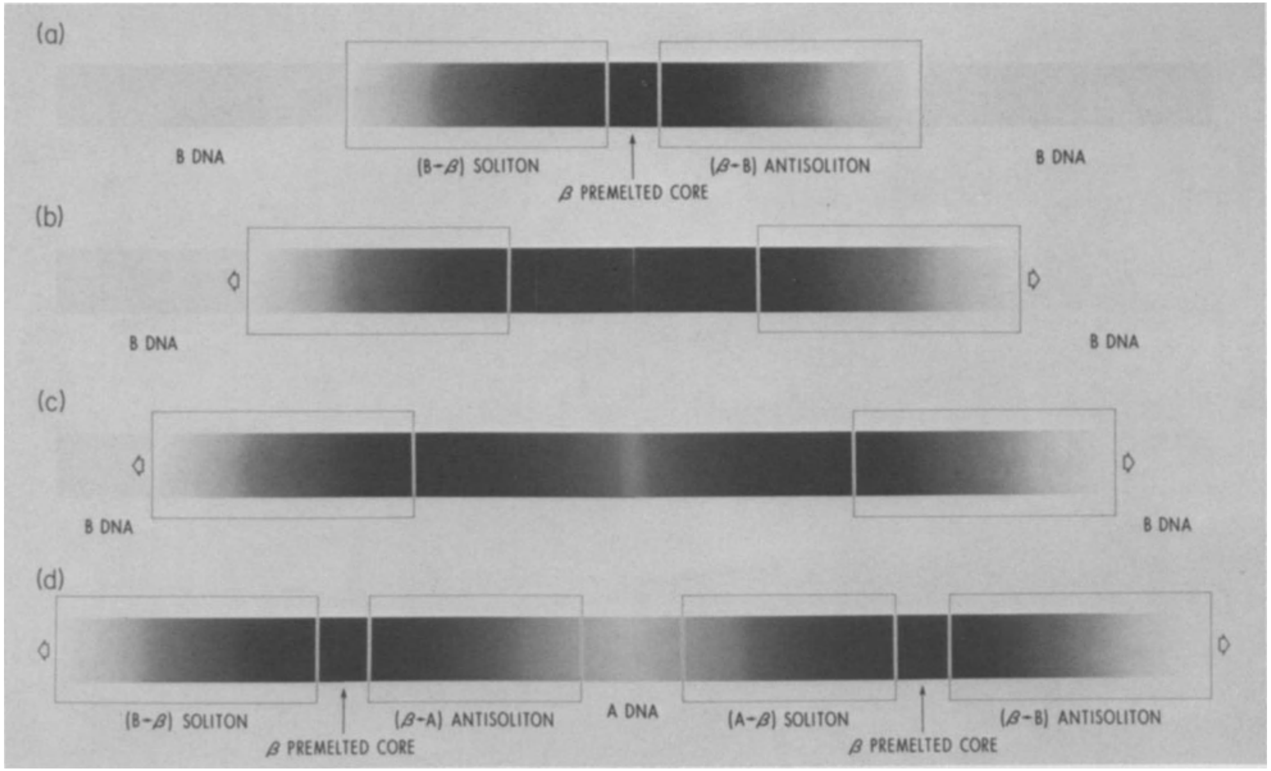


Figure 3. Schematic illustration of the B to A structural phase transition. (a) Nucleation begins at the centers of β premeltons within B DNA structure. (b) and (c) (B → β) solitons and (β → B) antisolitons move apart, leaving growing regions of β premelted cores whose centers begin to form A DNA. (d) Completion of the [(B → β):(β → A)] and [(A → β): (β → B)] β premelton pairs with intervening A DNA structure.

promoter by the RNA polymerase enzyme. RNA polymerase-promoter recognition could begin with the loose binding to β premelted nucleation sites within promoter regions, followed by a sequence of conformational transitions that progressively lead to a more open tight binding complex. Such an autocatalytic process can be viewed as one that creates "an avalanche of kinks", whose energies come from the stepwise interactions between the RNA polymerase and the promoter. This results in the formation and propagation of a new DNA phase -- such a phase could resemble the β structure, but be more completely unwound and single- stranded.

The process of RNA transcription could begin within this more open β premelton structure. The RNA polymerase enzyme then acts to catalyze the stepwise polymerization of ribonucleoside triphosphates into RNA, using one of two β denatured DNA chains as a template. Normal chain termination either occurs spontaneously, or requires other factors. Actinomycin and other related antibiotics are known to cause premature chain termination (22). We suggest that these antibiotics intercalate tightly into β premelted DNA regions within the transcriptional complex. This tends to immobilize the complex, increasing the probability of premature chain termination.

In recent years, increasing attention has focused on understanding the full range of conformational flexibility in DNA structure. This has culminated in the discovery of the Z DNA structure, a conformational state radically different from either B or A structures (23). The detailed nature of soliton excitations in Z DNA will be different from these described here. However, concepts of kink-antikink bound states as phase boundaries interconverting these forms will still be appropriate. We are attempting to model the molecular nature of the phase boundary connecting B with Z DNA with this in mind.

Acknowledgments:

This work has been supported in part by the National Institutes of Health and the Department of Energy. The paper has been assigned report No. UR-000-0000 at the DOE. Thanks also to D. Alexander and J.A. Krumhansel for providing stimulating discussions.

REFERENCES

1. A.S. Davydov, Physica Scripta, 20:387-394 (1979).
2. A.S. Davydov and N.I. Kislukha, Phys. Stat. Sol. 75B:735-742 (1976).
3. A.S. Davydov, Physca 3D 1:1-22, North Holland Publishing Company (1981).
4. S.W. Englander, N.R. Kallenbach, A.J. Heeger, J.A. Krumhansl and S. Litwin, Proc. Nat. Acad. Sci. USA 77,12:7222-7226 (1980).

5. A.C. Scott, F.Y.F. Chu and D.W. McLaughlin, *Proc. IEEE*, *61*:1443-1483 (1973).
6. A. Barone, F. Esposito, C.J. Magee and A.C. Scott, *Riv. Nuovo Cimento* *1*:227-267 (1971).
7. D. Alexander, J.A. Krumhansl, *in*: Structure and Dynamics: Nucleic Acids and Proteins, Adenine Press, Inc., pp. 61-80 (1983).
8. H.M. Sobell, E.D. Lozansky and M. Lessen, *Cold Spring Harb. Symp. Quant. Biol.* *43*:11-19 (1978).
9. H.M. Sobell, T.D. Sakore, S.C. Jain, A. Banerjee, K.K. Bhandary, B.S. Reddy and E.D. Lozansky, *Cold Spring Harb. Symp. Quant. Biol.*, *47*:293-314 (1983).
10. In previous communications, we have called this structure the "kink." However, the word "kink" has a broader meaning in the soliton physics area and, to avoid confusion, we have decided to rename this structure the β element. Its precise definition and full meaning is described in the text.
11. H.M. Sobell, B.S. Reddy, K.K. Bhandary, S.C. Jain, T.D. Sakore and T.P. Seshadri, *Cold Spring Harb. Symp. Quant. Biol.*, *42*:87-102 (1977).
12. The terms β DNA and β premelted DNA replace β kinked DNA used earlier.
13. The dip in the center of the energy density profile signifies the presence of a metastable structural state within the soliton-antisoliton bound state structure. This reflects the relaxation of strain energy in the sugar-puckering within the β premelted core region.
14. We use the term β premelton in the most general sense to describe kink-antikink bound states in double-helical DNA and RNA structure. Thus, premeltons can arise in B DNA and A DNA (or A RNA) and can act as phase boundaries transforming B DNA to A DNA during the B to A structural phase transition.
15. H.M. Sobell and S.C. Jain, *J. Mol. Biol.*, *68*:21-34 (1972).
16. G.J. Quigley, A.H.J. Wang, G. Ughetto, G. van der Marel, J.H. van Boom and A. Rich, *Proc. Nat. Acad. Sci. USA*, *77*:7204-7207 (1980).
17. R.J. Fiel and B.R. Munson, *Nucleic Acids Res.* *8*:2835-2842 (1980).
18. V.A. Bloomfield, D.M. Crothers and I. Tinoco, Jr., Physical Chemistry of Nucleic Acids, Harper and Row, Publishers (1974).
19. M.B. Fogel, S.E. Trullinger, A.R. Bishop and J.A. Krumhansl, *Phys. Rev. Lett.* *36*:1411-1414 (1976).
20. B. Jessee, G. Gargiulo, F. Razvi and A. Worcel, *Nucleic Acids Res.* *10*:*19*,5823-5834 (1982).
21. I.L. Cartwright and S.C.R. Elgin, *Nucleic Acids Res.* *10*:*19*, 5835-5852 (1982).
22. E. Reich and I.H. Goldberg, *in*: Progress in Nucleic Acid Research and Molecular Biology, *3*:183-234 (1964).
23. A.H.J. Wang, G.J. Quigley, F.J. Kolpak, G. van der Marel, J.H. van Boom and A. Rich, *Science*, *211*:*9*,171-176 (1981).

SOLITONS AND BIOENERGETICS

Alwyn C. Scott

Laboratory of Applied Mathematical Physics
The Technical University of Denmark
DK-2800 Lyngby, Denmark

INTRODUCTION

In recent theoretical[1] and numerical[2] studies of resonant energy transfer into a Davydov soliton on alpha-helix in protein it was assumed in a general way that hydrolysis of adenosine triphosphate (ATP) to adenosine diphosphate (ADP) would ultimately lead to soliton formation, thus:

$$\mathrm{ATP} + \mathrm{H_2O} \rightarrow \mathrm{ADP} + S_D + \text{other products} \tag{1}$$

where S_D denotes a Davydov soliton. The details of this reaction will be considered in the present paper. As a starting point for developing these details, we turn to the work of McClare[3-7] who has insisted that biological energy must be resonantly transferred into functional protein, but we give careful consideration to arguments by Davydov[8] that his soliton cannot directly absorb radiation. To resolve this difficulty we suggest an intermediate state, the proton soliton (S_P) which differs from the Davydov soliton because only the hydrogen bonded proton moves rather than an entire peptide unit. We emphasize that S_P is not a fictional entity: IR absorption by proton solitons in acetanilide has been observed. This is direct radiative transfer from the incident field into an S_P. Finally we posit the relaxation of a photon soliton into a true Davydov soliton.

The biochemical mechanism proposed here consists of four reactions, two of which are standard while the other two involve solitons. Since the soliton concept is rather new to biochemistry, we take extra care to establish the validity of our non-standard reactions.

RESONANT TRANSFER OF BIOLOGICAL ENERGY

The fundamental energy producing reaction in biological systems is generally accepted to be the hydrolysis of ATP; according to reference 5,

$$\boxed{ATP^{-4} + H_2O \rightarrow ADP^{-3} + HPO_4^{-2} + H^+ .} \tag{2}$$

Under normal physiological conditions this reaction produces about 0.49 ev of free energy[9]. McClare has shown that there are only two ways for this free energy to do useful work[3,4,7]: i) By a "constrained equilibrium" mechanism as in a conventional chemical engine (e.g. battery, internal combustion motor, etc.) in which free energy does useful work on a time scale that is long compared with that of thermal relaxation, and ii) A "molecular energy" mechanism in which free energy does useful work on a time scale that is short compared with the thermal relaxation time. Since constrained equilibrium mechanisms are necessarily macroscopic, they cannot be used for biological processes such as muscular contraction or active transport; thus we are forced to consider a molecular energy mechanism for the reaction shown in (2).

A molecular energy mechanism converts the free energy from a single reaction (2) into useful work. It cannot allow this energy to degrade into heat so it must act in a time short enough to avoid thermal relaxation, and this requirement strongly suggests resonant energy transfer. An obvious mechanism for resonant transfer is the relaxation of the proton produced in (2) back to water according to

$$\boxed{H^+ + OH^- \rightarrow H_2O^*} \tag{3}$$

where the asterisk denotes an excited state. On the right hand side of (3) the free energy is assumed to reside in vibrational quanta of the water molecule. At this point it is interesting to notice that the H-O-H bending mode of water has a frequency of 1646 cm^{-1} at 20°C[10] while from[1] the frequency of a soliton along a single channel of alpha-helix is

$$E_o - 2J = 1647 \ cm^{-1} . \tag{4}$$

Thus a Davydov soliton is almost resonant with the H-O-H bending mode of water and we are led to consider a third reaction[7]

$$H_2O^* \overset{?}{\rightarrow} S_D + H_2O . \tag{5}$$

Attractive as this reaction is, it must be rejected for reasons that have been discussed in detail by Davydov[8]. To appreciate this problem consider that the Davydov soliton is a localized state of the alpha-helix involving interaction between amide-I (essentially C=0 stretch) vibrations and longitudinal sound waves[8,11]. Localized amide-I bond energy acts to generate longitudinal sound which in turn reacts to trap the bond energy and prevent its dispersion. Creation of a soliton directly from radiation would require a time of the order of a period of longitudinal sound ($\sim 10^{-12}$sec) and during this time the frequency of the soliton state would be changing. Since resonance could not be maintained during the absorption process, it cannot take place. This "Frank-Condon principle" forbids direct radiative transfer of free energy from H_2O^* into a Davydov soliton as indicated in (5).

The restrictions of the Frank-Condon principle can be avoided for a soliton in which sound distortion takes place as fast as the amide-I vibrations. We consider such a soliton in the next section.

THE PROTON SOLITON

The Davydov soliton arises from a cooperative interaction between localized amide-I bond energy and lattice distortion. From reference 8 the interaction energy operator is

$$\hat{H}_{Di} = \chi_D \sum_{n,\alpha} (\hat{u}_{n+1,\alpha} - \hat{u}_{n\alpha}) \hat{B}^{\dagger}_{n\alpha} \hat{B}_{n\alpha} \tag{6}$$

where the subscript n counts turns of the helix, α (= 1, 2 or 3) specifies one of the three peptide units on each turn, $\hat{B}^{\dagger}_{n\alpha}(\hat{B}_{n\alpha})$ is a creation (annihilation) operator for amide-I vibrational quanta, and $\hat{u}_{n\alpha}$ is a longitudinal position operator for the (n,α) peptide unit. (See references 11 and 12 for figures that clarify the geometry.)

We are interested here in a soliton for which the time scale of the interaction energy is about equal to a period of amide-I vibration. Thus we choose the total interaction operator to be

$$\hat{H}_i = \hat{H}_{Di} + \hat{H}_{pi} \tag{7}$$

where

$$\hat{H}_{pi} \equiv \chi_p \sum_{n,\alpha} \hat{q}_{n\alpha} \hat{B}^{\dagger}_{n\alpha} \hat{B}_{n\alpha} \tag{8}$$

and $q_{n\alpha}$ is a position operator for displacement of the hydrogen bonded proton in the out-of-plane NH bending direction[13]. Then, following Davydov[8,11], we choose a soliton wave function

$$|\psi_{sol}\rangle = \sum_{n,\alpha} a_{n\alpha}(t)\, e^{\hat{\sigma}}\hat{B}^{\dagger}_{n\alpha}|0\rangle \tag{9}$$

where

$$\hat{\sigma} = -\frac{i}{\hbar}\sum_{n\alpha}[\beta_{n\alpha}(t)\,\hat{p}_{n\alpha} - \pi_{n\alpha}(t)\,\hat{u}_{n\alpha} + y_{n\alpha}(t)\,\hat{P}_{n\alpha} - \Pi_{n\alpha}(t)\hat{q}_{n\alpha}]. \tag{10}$$

Thus

$$\langle\psi_{sol}|\hat{H}_i|\psi_{sol}\rangle = \sum_{n\alpha}[\chi_D(\beta_{n+1,\alpha} - \beta_{n\alpha}) + \chi_p\, y_{n\alpha}]\,|a_{n\alpha}|^2 \tag{11}$$

where

$$\beta_n(t) = \langle\psi_{sol}|u_{n\alpha}|\psi_{sol}\rangle \tag{12}$$

is the average longitudinal displacement of the (n,α) peptide unit and

$$y_{n\alpha}(t) = \langle\psi_{sol}\,|\,\hat{q}_{n\alpha}|\psi_{sol}\rangle \tag{13}$$

is the average position of the corresponding hydrogen bonded proton.

Thus the probability amplitude for an amide-I vibrational quantum along the first chain, a_{n1}, is governed by the Schrödinger equation

$$\begin{aligned} i\hbar\,\dot{a}_{n1} &= [E_o + \chi_D(\beta_{n+1,1} - \beta_{n1}) + \chi_p\, y_{n1}]\, a_{n1} \\ &\quad - J(a_{n+1,1} + a_{n-1,1}) + L(a_{n2} + a_{n3}) \end{aligned} \tag{14}$$

where E_0 is the amide-I vibrational energy and J and L are respectively longitudinal and transverse dipole-dipole interaction energies. Also we have the equation

$$m_o\ddot{\beta}_{n1} - w(\beta_{n+1,1} - 2\beta_{n1} + \beta_{n-1,1}) = \chi_D(|a_{n1}|^2 - |a_{n-1,1}|^2) \tag{15}$$

which describes longitudinal motion of peptide units along the first chain. Dynamics of the hydrogen bonded protons are governed by

$$m_p(\ddot{y}_{n1} + \omega_p^2 y_{n1}) = -\chi_p|a_{n1}|^2 \tag{16}$$

where m_p is the proton mass and ω_p is the frequency of out-of-plane NH bending ($700\ \mathrm{cm}^{-1}$)[13]. The time scale for proton motion is $\omega_p^{-1} = 7.6 \times 10^{-15}$ sec while that for the lattice motion is $[m_o/w]^{\frac{1}{2}} = 1.21 \times 10^{-13}$ sec. Thus for times within the range

$$[m_o/w]^{\frac{1}{2}} > t > \omega_p^{-1} \tag{17}$$

we expect motion of the protons without longitudinal displacements of the peptide units. Since the time ω_p^{-1} is approximately equal to a period of amide-I vibration, the Frank-Condon principle does not forbid the formation of a soliton which involves interaction of amide-I bond energy with proton displacement. We call this a "proton soliton" to avoid confusion with the Davydov soliton which may develop when

$$t > [m_o/w]^{\frac{1}{2}} . \tag{18}$$

For the time range (17)

$$y_{n\alpha} \doteq -\frac{\chi_p}{m_p \omega_p^2} |a_{n\alpha}|^2 \tag{19}$$

while $\beta_n \doteq 0$ so (14) can be written

$$i\hbar\, \dot{a}_{n1} - (E_o - 2J)a_{n1} + J(a_{n+1,1} - 2a_{n1} + a_{n-1,1}) + \frac{\chi_p^2}{m_p \omega_p^2} |a_{n1}|^2\, a_{n1} \doteq L(a_{n2} + a_{n3}) . \tag{20}$$

We now assume that the number of helix turns spanned by a soliton is large enough to permit the continuum approximation

$$\frac{a_{n+1,\alpha} - 2a_{n\alpha} + a_{n-1,\alpha}}{b^2} = \frac{\partial^2 a_\alpha}{\partial z^2} \tag{21}$$

where $z = nb$ and b is the pitch of the helix. Then scaling the space variable by $z = b(J/\hbar)^{\frac{1}{2}}x$ and the amplitudes by $a_\alpha = (\hbar/J)^{\frac{1}{4}}A_\alpha$ puts (20) into the form

$$i\dot{A} - \hbar^{-1}(E_o - 2J)\, A_\alpha + A_{\alpha,xx} + \gamma_p |A_\alpha|^2 A_\alpha \doteq \hbar^{-1} L(A_{\alpha+1} + A_{\alpha-1}) \tag{22}$$

where

$$\gamma_p \equiv \frac{\chi_p^2}{m_p \omega_p^2 \sqrt{hJ}} . \tag{23}$$

Then we can define $\bar{A} = \mathrm{col}(A_1, A_2, A_3)$ and obtain both for symmetric solutions $(A_1 = A_2 = A_3)$ and antisymmetric solutions $(A_1 + A_2 + A_3 = 0$ but $|A_1| = |A_2| = |A_3|)$ the vector nonlinear Schrödinger equation[12]

$$i\dot{\bar{A}} - \hbar^{-1}(E_o - 2J)\bar{A} + \bar{A}_{xx} + \frac{\gamma_p}{3}(\bar{A}^* \cdot \bar{A})\bar{A} = \hbar^{-1} Z\, \bar{A} \tag{24}$$

where

$$Z \equiv \begin{vmatrix} 0 & L & L \\ L & 0 & L \\ L & L & 0 \end{vmatrix} \quad . \tag{25}$$

For a single quantum of amide-I vibrational energy a stationary symmetric proton soliton can be written

$$\bar{A} = \mathrm{col}(A_s, A_s, A_s) \tag{26a}$$

where

$$A_s = \frac{\sqrt{2\gamma_p}}{12} \operatorname{sech}\left(\frac{\gamma_p}{12}\, x\right) \exp\left[\, i\, \hbar^{-1}\left(-E_o + 2J - 2L - \frac{\gamma_p^2}{144}\right) t \,\right] . \tag{26b}$$

There is also a stationary antisymmetric proton soliton which can be written

$$\bar{A}_a = \mathrm{col}\left(A_a,\ A_a\, e^{2\pi i/3},\ A_a\, e^{-2\pi i/3}\right) \tag{27a}$$

where

$$A_a = \frac{\sqrt{2\gamma_p}}{12} \operatorname{sech}\left(\frac{\gamma_p}{12}\, x\right) \exp\left[\, i\, \hbar^{-1}\left(-E_o + 2J + L + \frac{\gamma_p^2}{144}\right) t \,\right] . \tag{27b}$$

The threshold conditions for establishing a proton soliton can be estimated from reference 1 by replacing the longitudinal spring constant $w = 13$ newtons per meter with $m_p\omega_p^2 = 29$ newtons per meter for a proton soliton. Thus the "resonance" level of exciton-proton coupling at which two quanta of input are expected to launch a single quantum proton soliton at one end of a chain is

$$\chi_{pr} = 8.3 \times 10^{-11} \text{ newtons} \ . \tag{28}$$

Infrared absorption by proton solitons has been observed experimenttally in acetanilide[14]. Assuming the 15 cm^{-1} red shift of the 1650cm^{-1} band to be binding energy for a proton soliton, the exciton-proton coupling constant is calculated as [14]

$$\chi_p \doteq 20 \times 10^{-11} \text{ newtons} . \qquad (29)$$

Thus we expect the reaction indicated in (3) can be followed by the transition of excited water (H_2O^*) into one or perhaps two proton soliton(s) according to

$$\boxed{H_2O^* \rightarrow S_p + H_2O .} \qquad (30)$$

On the left hand side it is assumed that the excited water carries two quanta of the H-O-H bending vibration (.41 ev) and that at least one of these quanta generates a proton soliton.

THE DAVYDOV SOLITON

On the time scale indicated in (18) we assume that reaction (30) has proceeded and a proton soliton has formed. Returning to (14) we no longer assume $\beta_{n\alpha}(t)$ to be zero; its time evolution is now governed by (15). We have shown by analytical arguments in reference 1 (and confirmed by numerical calculations in reference 2) that initialization with two amide-I quanta localized over four adjacent bonds will excite a single quantum Davydov soliton when the resonance value of the excition phonon coupling parameter is

$$\chi_{Dr} = 4.7 \times 10^{-11} \text{ newtons} . \qquad (31)$$

This compares favorably with the self-consistent field calculations of Kuprievich and Kudritskaya[15] which indicate that

$$\chi_D = 3 - 5 \times 10^{-11} \text{ newtons} . \qquad (32)$$

In the present case, however, the threshold situation is somewhat different: we assume that a single quantum proton soliton is initializing the system defined by (14) and (15). From (11) we see that the total exciton-phonon interaction energy will become more negative when the effect of longitudinal motion of the peptides is considered during the time scale indicated in (18). Thus the binding energy of a Davydov soliton solution for the system (14) - (15) will always be greater than that for a proton soliton alone, and it is energetically favorable for a proton soliton to decay into a Davydov soliton (D_D) according to

$$\boxed{S_p \rightarrow S_D + \text{phonons}\ .} \tag{33}$$

The phonons involved in (33) have been discussed theoretically and described numerically in reference 16. They travel at the speed of longitudinal sound which is $[w/m_0]^{\frac{1}{2}}$ helical turns per second or one helical turn in about .12 picosecond. The relaxation time (τ) required in (33) is therefore

$$\tau = n[m_o/w]^{\frac{1}{2}} \tag{34}$$

where n is the width of a Davydov soliton measured in turns of the helix. From II we found $n \sim 5$ so the relaxation time in (33) is about

$$\doteq 0.6 \times 10^{-12} \text{ seconds}\ . \tag{35}$$

The half width ($\Delta\omega$) of the proton soliton absorption is about 7 cm^{-1} which implies[14]

$$\frac{1}{\Delta\omega} = .76 \times 10^{-12} \text{ seconds}\ . \tag{36}$$

The approximate agreement between the relaxation times in (35) and (36) suggests that the proton soliton produced by infrared absorption in acetanalide decays into a Davydov soliton according to (33).

SUMMARY OF THE PROPOSED MECHANISM

The mechanism that we propose for the transfer of free energy from ATP hydrolysis into a protein is embodied in the reactions (2), (3), (30), and (33) which are rewritten here for the convenience of the reader.

$$ATP^{-4} + H_2O \rightarrow ADP^{-3} + HPO_4^{-2} + H^+(0.49) \tag{A}$$

$$H^+(0.49) + OH^- \rightarrow H_2O^*(0.41) \tag{B}$$

$$H_2O^*(0.41) \rightarrow (1 \text{ or } 2)\ S_p(0.205) + H_2O \tag{C}$$

$$S_p(0.205) \rightarrow S_D(0.205) + \text{phonons}\ . \tag{D}$$

Although, as we frankly admit, this reaction mechanism is speculative, it has the advantage of being quite specific and therefore subject to the rigors of experimental investigation. In particular we can say just how much free energy is preserved at each step and where it is located. For example:

A) In the first reaction a free energy of 0.49 ev is carried by the extra hydrogen ion (H^+) produced during ATP hydrolysis.

B) In the second reaction this hydrogen ion combines with OH^- to produce an excited water molecule that carries two vibrational quanta (0.41 ev) of free energy in H-O-H bending.

C) In the third reaction the free energy of at least one (possibly both of these H-O-H vibrational quanta is transferred to at least one (possibly two) proton solitons.

D) The free energy of a proton soliton (0.205 ev) is transferred directly to a Davydov soliton.

Thus this mechanism has an efficiency of 42% if only one proton soliton is produced in the third reaction. If two proton solitons are produced, the efficiency will be 84%,

In any resonant energy transfer mechanism involving water one is forced to consider whether the water will not also act in an undesired way to dissipate free energy. If water can send energy in resonantly, why can't it carry energy away just as easily? McClare[3-7] has dealt with this question by assuming that the energetic states of a biomolecule could be symmetric (with a large dipole moment) for the absorption of radiation and antisymmetric (with a much smaller dipole moment) for the preservation of free energy. As we see from (26) and (27), solitons on alpha-helix come in both symmetric and antisymmetric configurations and therefore fully satisfy McClare's requirements. Recent electromagnetic calculations indicate that the lifetime of a symmetric soliton on alpha-helix that is totally immersed in water is about 500 picoseconds while the lifetime of an antisymmetric soliton must be much longer[17]. In the present picture water losses from a Davydov soliton should be somewhat lower since it will be tuned 10 - 20 cm^{-1} below a proton soliton because of its extra binding energy. Finally we expect the Frank-Condon principle[8] to be just as effective in preventing direct radiation from Davydov soliton as in preventing direct radiative absorption.

REFERENCES

1. Scott, A.C. (Physica Scripta to appear).
2. MacNeil, L. and Scott, A.C. (Physica Scripta to appear).
3. McClare, C.W.F., J. Theor. Biol. 30, 1 (1971).
4. McClare, C.W.F., Nature 240, 88 (1972).
5. McClare, C.W.F., J. Theor. Biol. 35, 233 (1972).
6. McClare, C.W.F., J. Theor. Biol. 35, 569 (1972).
7. McClare, C.W.F., Ann. N.Y. Acad. Sci. 227, 74 (1974).
8. Davydov, A.S., Biology and Quantum Mechanics, Pergamon Press (1982).

9. Hock, F.L., Energy Transformations in Mammals: Regulatory Mechanisms, Saunders, Philadelphia (1971).
10. Fox, J.J. and Martin, A.E., Proc. Roy. Soc. (London) A 174, 234 (1940).
11. Davydov, A.S., Physica Scripta 20, 387 (1979).
12. Scott, A.C., Physica Scripta 25, 651 (1982).
13. Abbott, N.B. and Elliott, A., Proc. Roy. Soc. (London) A 234, 247 (1956).
14. Careri, G., U. Buontempo, F. Carta, E. Gratton and A.C. Scott, (Phys. Rev. Letts. to appear).
15. Kuprievich, Y.A. and Kudritskaya, Z.G. Preprint ITP-82-64E, Institute for Theoretical Physics, Kiev (1982).
16. Scott, A.C. Phys. Rev. A 26, 578 (1982); ibid. 27, 2767 (1983).
17. Scott, A.C., Phys. Letts. A 94, 193 (1983).

NONLINEAR DYNAMICS OF GLOBULAR PROTEINS

Peter S. Lomdahl

Center for Nonlinear Studies
Los Alamos National Laboratory
Los Alamos, NM 87545

1. INTRODUCTION

The general opinion and knowledge about nonlinear phenomena in the physical sciences has grown quite remarkably within the last 20 years. The prevalence of nonlinear processes in nature has been recognized for a long time - but until recently - also often neglected when dealing with specific problems. This approach was necessary in order to get tractable models which could be handled with the well proven linear mathematical tools, such as Fourier analysis and superposition principles. The fruitful collaboration over the last two decades between applied mathematicians, physicists, and engineers - has however brought about new mathematical tools which makes it possible to deal with certain nonlinear problems in a more systematic way. In particular, there are systems in which nonlinear and dispersive effects coexist and compete - this competition is often relieved by formation of a stable pulselike object - it is in such systems that the concept of the soliton plays a useful role. Especially when viewed as a paradigm for a new nonlinear "normal" mode. This approach has already been applied with success in a number of areas of physics where dispersive and nonlinear effects are important, i.e. condensed matter, plasmas, and optics.[1] So it is with some optimism that we are now facing the similar but much more complicated problems in biological systems.

DAVYDOV has as a first step in this direction considered a regular chain of hydrogen-bonded peptide groups. I shall not outline the details of the theory here - it is well described in the references[2,3] and also in A. C. SCOTT's contribution to these proceedings. Also, the mathematical formalism is in certain cases identical to that of the molecular crystal model developed by HOLSTEIN.[4] However, I would like to emphasize the following basic rationals behind the theory: i) All peptide groups have a particularly characteristic intramolecular vibration called amide-I, consisting mainly of C = 0 valence vibrations with a frequency of 1650 cm^{-1} (~0.21 eV). It is tempting to view the amide-I mode as a possible "basket" for storage of energy (released during ATP hydrolysis (0.43 eV ~ 2 × 0.21 eV) e.g.). ii) One mechanism which tends to act against localization of amide-I energy is due to resonant transfer of energy between the identical peptide groups in the molecular chain. This dispersion mechanism can be approximated as an electrostatic dipole-dipole interaction and is particularly important because the transition dipole moment of the amide-I vibration is relatively large (d ~ 0.3D). iii) Counteracting this mechanism, is the coupling between the amide-I vibration and other intra- and intermolecular excitations. In DAVYDOVs original work, emphasis was put on the displacement of the peptide group as a whole. When a particular peptide group is excited with amide-I energy its interaction with neighboring peptide groups changes, resulting in a change in the equilibrium distance between groups. The actual energy of a peptide group thus depends on its displacement, making the system nonlinear. In the recent work by CARERI et al.[5] the displacement of the proton intrinsic to the peptide group was considered as an intramolecular excitation coupled to the amide-I vibration. Here, the result is essentially the same - the coupling leads to focusing of amide-I bond energy.

The above mentioned competition between dispersion and focusing of amide-I energy, will in a regular system of hydrogen-bonded peptide groups - like α-helix proteins or acetanilide crystals - lead to the formation of a soliton like object which can travel along the chain without changing their shape. This is essentially a manifestation of the fact that the system has perfect translational symmetry. A natural question to ask is: "How is this competition relieved in a real globular protein?" Such a system will in general not have translational invariance between the different peptide groups, and soliton formation is not to be expected. However, the mechanisms for dispersion and focusing of amide-I energy will still be present.

In this paper I will describe some ongoing work which is aimed at generalizing DAVYDOVs ideas to a real globular protein. So far, a computer code, "GLOP", which calculates amide-I bond energy evolution on a globular protein has been developed and tested. The code is quite versatile and take as input the coordinates of a protein. The full geometry of the molecule is then taken into account when the dipole-dipole interaction between peptide groups is calculated. The amide-I energy is coupled to one intramolecular excitation, but can without difficulty be extended to more or to include intermolecular excitations. The outline of the paper is as follows: In section 2, I describe the model Hamiltonian, in section 3 preliminary results for two structures are given, the first - acetanilide is a test case for a regular chain structure of peptide groups, the second is for the globular protein - lysozyme. Finally in section 4, I comment on future work and relations to similar work in other fields.

2. MODEL HAMILTONIAN

In this section, I will derive the equations of motion for amide-I bond energy evolution on a protein structure which does not necessarily have a regular symmetry. This set of equations form the basis for the construction of the numerical code. The analysis is with a few exceptions, quite in the spirit of DAVYDOV[2,3].

The amide-I vibrational excitation will be treated as a one-excited state in terms of quantum mechanics. It will be coupled to an intramolecular excitation which will be treated as a classical variable from the beginning. Here I specifically consider the out-of-plane vibration of the proton of each peptide group[5] and assume that this motion can be described quasiclassically as an Einstein oscillator. I should stress that this simplified approach does not put limitations on the final code - the coupling to other intra- or intermolecular excitation can easily be included.

If $b_j^\dagger$ and b_j are boson creation and annihilation operators for an amide-I excitation on the j-th peptide group, then the Hamiltonian for the pure amide-I system is given by

$$H_A = \sum_j \varepsilon_j b_j^\dagger b_j + \sum_{j \neq k} \sum_k \frac{V_{jk}}{2} (b_k^\dagger b_j + b_j b_k^\dagger) \, . \tag{1}$$

The summation is carried out over all N peptide groups in the system. The first term in Eq. (1) defines the amide-I excitation energy without considering the resonance dipole-dipole interaction. Here ε_j is the energy of the amide-I vibration on the j-th peptide group. This energy varies slightly from group to group - depending upon the actual configuration in which the peptide group is located (α-helix, β-sheet, random coil, etc.). For a regular chain structure it is of course a constant. The operators in the last term in Eq. (1) represent transfer of amide-I energy from group j to k (and vice versa) due to dipole-dipole interaction. The quantity V_{jk} is the dipole-dipole interaction energy between groups j and k.

This energy is calculated in the usual electrostatic approximation. Following CHIRGADZE and NEVSKAYA[6] the location of the center of the transition dipole ($\vec{d}_j = d\vec{r}_j$) is placed about 0.4 Å from the oxygen atom in the oxygen-nitrogen direction (see Fig. 1).

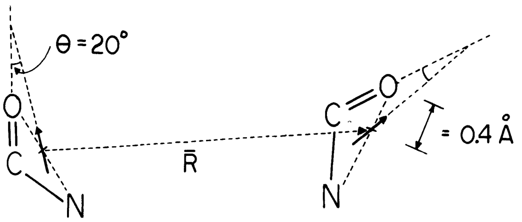


Fig. 1. Dipole-dipole interaction between peptide groups. A unit vector in the $\bar{R}$ direction is $\bar{n} = \bar{R}/|\bar{R}|$.

The direction of the transition dipole moment is taken as lying in the amide group plane and making an angle of 20° between the C = O bond, being inclined towards the direction of the CN bond. This approach is what was used by MOORE and KRIMM in their analysis of polyglycine.[7] Knowing the location - the interaction energy is easily calculated. Dipole k sees the potential: $U(\bar{R}) = d\bar{r}_j\cdot\bar{R}/(4\pi\varepsilon R^3)$. The field is: $\bar{E} = -\nabla_{\bar{R}} U(\bar{R})$ or

$$\bar{E} = -\frac{d}{4\pi\varepsilon}\left[\frac{\bar{r}_j}{R^3} - \frac{3(\bar{r}_j\cdot\bar{R})}{R^4}\bar{n}\right]$$

$$= -\frac{d}{4\pi\varepsilon R^3}\left[\bar{r}_j - 3(\bar{r}_j\cdot\bar{n})\bar{n}\right] .$$

The interaction energy is thus: $V_{jk} = -\bar{E}\cdot d\bar{r}_k = V_{kj}$ or

$$V_{jk} = \frac{d^2}{4\pi\varepsilon R^3}\left[\bar{r}_j\cdot\bar{r}_k - 3(\bar{r}_j\cdot\bar{n})(\bar{r}_k\cdot\bar{n})\right] . \tag{2}$$

The dielectric constant, ε, is taken as the screened value: $\varepsilon(R) = \varepsilon_o$ for $R \leq R_o$ and $\varepsilon(R) = n\varepsilon_o$ for $R > R_o$. This reflects the fact that as $R \to 0$ there is no intervening matter between atoms to increase ε. Thus $\varepsilon = \varepsilon_o$ for R sufficiently small. In the code R_o and n are adjustable parameters. For the results reported in section 3, R_o and n were taken as 15 Å and 1.6, respectively.

The next term we consider in our description of the peptide group system is the energy associated with the interaction of the amide-I excitation and the vibrations of the hydrogen-bonded proton in the out-of-plane NH bending direction. If y_j is a coordinate denoting the displacement away from equilibrium of the proton on the j-th peptide group, then the interaction energy is

$$H_I = \chi \sum_j y_j b_j^\dagger b_j \ . \tag{3}$$

It reflects the fact that the amide-I energy depends on the distortion of the adjacent proton. The coupling constant, χ, is in fact the change in amide-I energy per unit extension of the NH bond.

The motion of the proton itself is governed by the classical Hamiltonian:

$$H_p = T + U = \frac{m_p}{2} \sum_j \dot{y}_j^2 + \frac{m_p \omega_p^2}{2} \sum_j y_j^2 \ . \tag{4}$$

The mass of the proton is m_p, while ω_p denotes the frequency of out-of-plane NH bending (~700 cm^{-1}).[8,5]. The first term is the kinetic energy and the second the potential energy.

The total Hamiltonian $H = H_A + H_p + H_I$ of the system must satisfy the time dependent Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle = (H_A + H_p + H_I)|\psi\rangle \ . \tag{5}$$

The state of the system is described by the wavefunction, $|\psi\rangle$, which is taken to have the form

$$|\psi\rangle = \sum_j \tilde{a}_j(t) b_j^\dagger |0\rangle \ , \tag{6}$$

where the $\tilde{a}_j(t)$ are complex numbers whose values at some initial time are arbitrary (apart from normalization), and where the ket, $|0\rangle$, denotes the "vacuum" state of the amide-I excitation, also defined as the state for which $b_j|0\rangle = 0$. For simplicity the normalization is choosen as:

$$\langle\psi|\psi\rangle = \sum_j |\tilde{a}_j(t)|^2 = 1 \ . \tag{7}$$

Inserting Eqs. (1), (3), and (4) in Eq. (5) and using Eqs. (6) and (7) the following equations are obtained:

$$i\hbar \dot{\tilde{a}}_j = [\varepsilon_j + T + U]\tilde{a}_j + \chi y_j \tilde{a}_j + \sum_{k \neq j} V_{jk} \tilde{a}_k \ . \tag{8}$$

In this set of equations j runs from 1 to N (= number of peptide groups). As a closure, a set of equations for y_j is needed, this is readily obtained from H by applying Hamiltons equations ($\dot{q} = \frac{\partial H}{\partial p}$, $\dot{p} = -\frac{\partial H}{\partial q}$, here $q = y_j$, and $p = m_p\dot{y}_j$), combined with Eqs. (6)-(7):

$$\ddot{y}_j + \omega_p^2 y_j = -\frac{\chi}{m_p}|\tilde{a}_j|^2 \; . \tag{9}$$

The Eqs. (8)-(9) are the main result of the analysis and form the basis for the code GLOP. They describe the time evolution of amide-I bond energy coupled to vibrations of the proton in the peptide group. The quantity $|\tilde{a}_j(t)|^2$ characterizes the distribution of amide-I energy over the individual peptide groups of the protein.

To get a better understanding for the behavior of Eqs. (8)-(9), it is instructive to look at time scales which are slow compared to ω_p^{-1}. In this range $\ddot{y}_j \cong 0$ and Eqs. (8) and (9) can be combined to:

$$i\hbar\,\dot{\tilde{a}}_j = [\varepsilon_j + U]\tilde{a}_j - \frac{\chi^2}{m_p\omega_p^2}|a_j|^2 a_j + \sum_{k\neq j} V_{jk}\tilde{a}_k \; . \tag{10}$$

Further, introducing $a_j(t) = \tilde{a}_j(t)e^{-iUt/\hbar}$ - together with the scaled time variable $\tau = t/T_o$, where $T_o = \hbar\omega_p^2 m_p/\chi^2$ - we get:

$$i\,\frac{da_j}{d\tau} = \delta_j a_j - |a_j|^2 a_j + \sum_{k\neq j} \gamma_{jk} a_k \; , \tag{11}$$

here $\delta_j = T_o\varepsilon_j/\hbar$ and $\gamma_{jk} = T_o V_{jk}/\hbar$. Eq. (11) can also be written in matrix form:

$$i\,\frac{d\bar{A}}{d\tau} = \bar{\bar{M}}\bar{A} \; , \tag{12}$$

where $\bar{A} \equiv \mathrm{col}[a_1, a_2, \ldots, a_N]$ and

$$\bar{\bar{M}} \equiv \begin{bmatrix} \delta_1-|a_1|^2 & \gamma_{12} & \gamma_{13} & \cdot & \cdot & \gamma_{1N} \\ \gamma_{12} & \delta_2-|a_2|^2 & \gamma_{23} & \cdot & \cdot & \gamma_{2N} \\ \gamma_{13} & \gamma_{23} & \delta_3-|a_3|^2 & \cdot & \cdot & \gamma_{3N} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \gamma_{1N} & \gamma_{2N} & \cdot & \cdot & \cdot & \delta_N-|a_N|^2 \end{bmatrix}_{N\times N} \tag{13}$$

In the form of Eqs. (12)-(13), it is readily seen that diagonal terms in $\bar{\bar{M}}$ tend to focus amide-I energy while the off-diagonal terms are responsible for dispersion of energy. For numerical purposes it is useful to note the following conservation law for Eq. (12). Multiplying Eq. (12) by $i\bar{A}^*$ one gets: $\dot{\bar{A}}\cdot\bar{A}^* = -i\bar{\bar{M}}|\bar{A}|^2$, but since $\bar{\bar{M}}$ is real we also have: $\dot{\bar{A}}^*\bar{A} = i\bar{\bar{M}}|\bar{A}|^2$, from which it follows that $\dot{\bar{A}}\bar{A}^* + \dot{\bar{A}}^*\bar{A} = 0$ or $\bar{A}\bar{A}^* =$ constant. Considering Eq. (7) this is of course not surprising, but it is useful in monitoring the numerical integration of Eq. (12).

The code GLOP has been set up to integrate either Eqs. (8)-(9) or Eq. (11) with a given initial configuration of amide-I bond energy, $a_j(t=0)$, - and for Eqs. (8)-(9) initial proton displacement, $y_j(t=0)$. As a start the coordinates for the desired protein structure are input and the dipole-dipole interaction energies, V_{jk}, are calculated from Eq. (2). The equations of motion are then integrated using a versatile leap-frog predictor-corrector method[9]. The accuracy of the integration is monitored at each time-step via the conserved quantity. In addition to instantanous pictures of $|a_j(t)|^2$ during the time evolution, quantities measuring total "bending" - like $|\sum_j |a_j|a_j|$ - are recorded as a function of time. Various diagnostics - including Fourier analysis - can then be applied to such time series.

3. EXAMPLES

In this section, I will present two cases of what can be obtained from actual numerical integration of Eq. (11). These are only examples of preliminary results with the code GLOP - a detailed study of both systems will be published elsewhere.

In order to test the code I have first studied a regular lattice system - acetanilide (or ACN) which has recently been investigated quite extensively, both theoretically and experimentally[5]. The system basically consists of two close chains of hydrogen-bonded peptide groups. Attached to each group is a benzene ring and a methyl group[10]. In Fig. 2, two "unit cells" are shown. The dark atoms belong to one chain, while the white ones belong to the other. Some atoms in the dark chain are partly covered by those of the white chain. The actual system considered consisted of 50 "unit cells". The dipole-dipole interaction in this system will be quite regular because of the translational invariance of the lattice. In Fig. 3 the number of peptide groups with a particular interaction energy is shown as a function of energy. It is seen that only a few destinct values play an important role. In fact, if

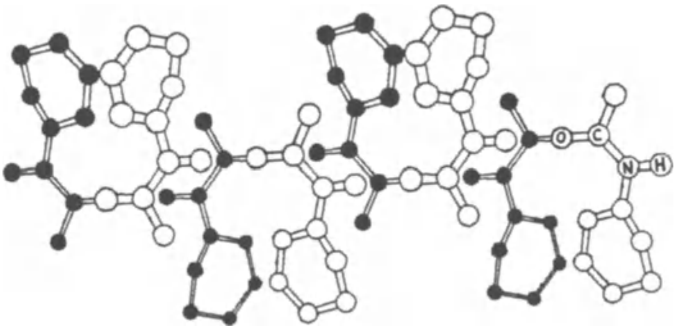


Fig. 2. Acetanilide (or ACN) consisting of two close chains of peptide groups. Two "unit cells" are shown.

only nearest neighbor interaction between groups is taken into account, the system can be described by two coupled nonlinear Schrodinger equations[5], and is thus expected to support soliton-like excitations.

This is exactly what is depicted in Fig. 4. It shows the time evolution of $|a_j(t)|^2$ resulting from the numerical integration of Eq. (11). The results are displayed for 50 different times in the interval $0 < t \lesssim 60$ ps. The peptide groups are numbered from 1 to 100, where 1-50 correspond to groups on one chain and 51-100 to groups on the other chain. The initial condition was one quantum of amide-I energy placed on group number one. A non-dispersive soliton-like object is seen to form, and propagate down the chain system. In addition, it is seen that the energy is oscillating between the two chains. When $|a_j(t)|^2$ is large on one chain it is

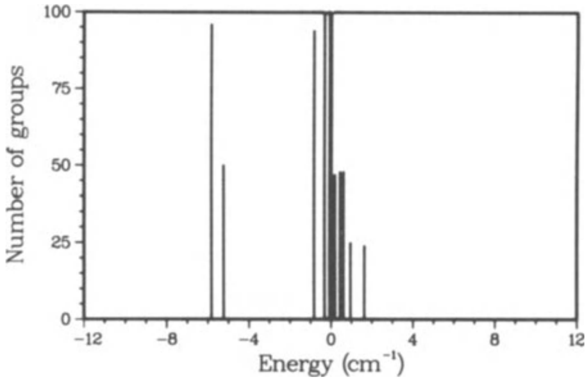


Fig. 3. Dipole-dipole interaction energy distribution for acetanilide.

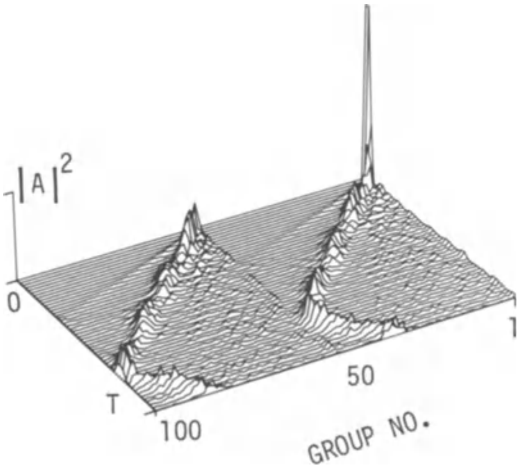


Fig. 4. Result of numerical solution of Eq. (11) with one quantum of amide-I energy placed at peptide group no. 1 in the acetanilide system. Groups with no. 1-50 are on one chain and groups with no. 51-100 on the other chain. The time evolution is shown for the first 60 ps approx. The coupling constant, χ, was $1.0 \cdot 10^{-10}$ N (see Ref. 5).

low on the other. This behavior is to be expected from an analysis based on two coupled nonlinear Schrödinger equations[5]. The result of Fig. 4 is confirming that this is actually a good approximation - even when the full dipole-dipole interaction is taken into account the nearest neighbor interaction tells the main story.

To test the code on a real globular protein - preliminary runs with the enzyme lysozyme (from chicken egg-white) have been done. The structure is well known[11] and consists of 129 peptide groups. The carbon atoms in the peptide groups are shown connected as a "backbone" in Fig. 5. The overall non-regular structure is quite clear. Part of the structure is however in the α-helix configuration, those groups (5-15, 24-34, 88-96) are indicated by black carbon atoms in Fig. 5. The number of groups with a particular dipole-dipole interaction energy is shown in Fig. 6. The disordered structure is clear when compared with Fig. 3 for ACN. In Fig. 7 we show the result of the solution of Eq. (11) with initially two quanta placed on peptide groups 33 and 34 which is the end of the longest α-helix part. It is interesting to note that the energy basically stays localized on the α-helix part - it oscillates between the two ends for approximately 225 ps (as far as the integration was performed). Additional runs where the energy was initiated on other peptide groups (non α-helical), showed that dispersion of energy easily can occur too.

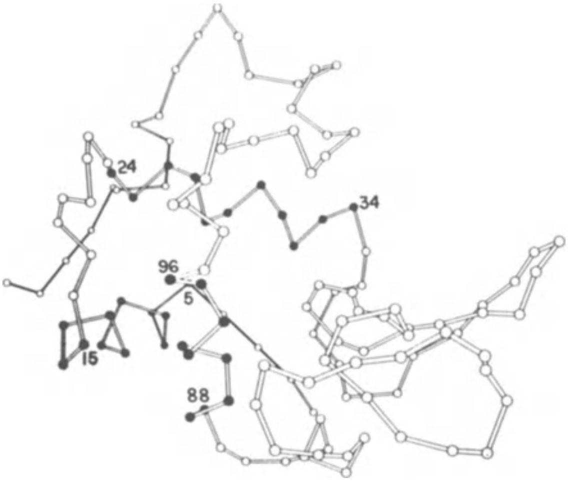


Fig. 5. Chicken egg-white lysozyme. The carbon atoms (from the C = 0 group, not α-carbon) in each peptide group are shown. The black atoms indicate α-helix regions (peptide groups 5-15, 24-34, 88-96).

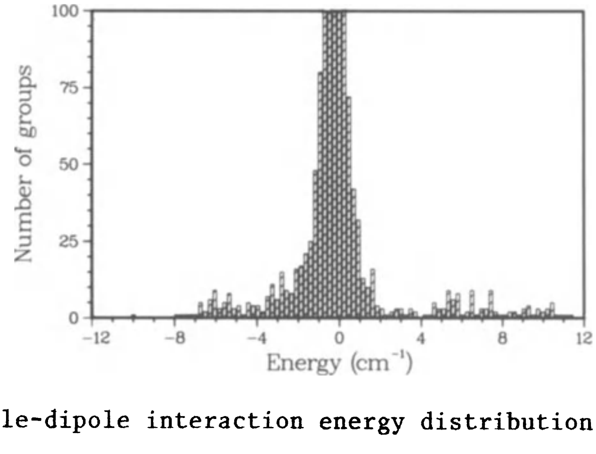


Fig. 6. Dipole-dipole interaction energy distribution for lysozyme.

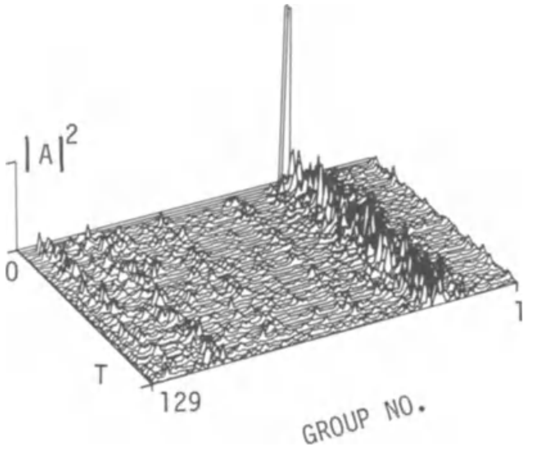


Fig. 7. Results of numerical solution of Eq. (11) with two quanta of amide-I energy placed at peptide groups no. 33 and 34 (α-helix). The time evolution is shown for approx. 225 ps. The coupling constant, χ, was chosen to $0.4 \cdot 10^{-11}$ N and $m_p \omega_p^2$ was equal to 13 N/m.

4. CONCLUDING REMARKS

The examples given above are of course not exhaustive and more work is needed before conclusive statements can be made. In addition to the study of the specific systems mentioned above, work is in progress to improve the code. In particular the following areas are considered.

1) Finding localized stationary states. The possible initial conditions used for the dynamics calculations with Eq. (11) are not ideal at the present time. Stationary states - i.e. solutions to Eq. (12) of the form $\bar{A} = \bar{A}_o e^{-i\omega t}$, where ω is a constant - would form a better starting point. Note that the superposition of such states is in general not a new stationary state. The nonlinear dynamics of Eq. (11) would govern the evolution. Inserting $\bar{A} = \bar{A}_o e^{-i\omega t}$ in Eq. (13) leads to a nonlinear eigenvalue problem, where Eq. (7) is a constraint equation. The solution of this problem is the main difficulty in finding the stationary states. Of particular interest are those states which are localized and have low energy (small ω). Such states could be expected to have soliton-like character if they were localized in places where the protein structure was fairly regular.

In connection with this problem it is interesting to note that related work has been going on for some years in the area of amorphous semiconductors. Here equations similar to Eq. (11) arise naturally too. We refer the interested reader to the review by THOULESS[11].

2) More realistic coupling to other intra- and intermolecular modes. The simple system represented by Eq. (9) should be extended to include more intra- and intermolecular excitations. More conventional molecular dynamics codes[12] are already considering such systems and should be incorporated in GLOP. In this respect it is also important to consider dissipative effects (energy transfer from the protein into water e.g.). As a starter this could be done by including phenomenological damping terms in Eq. (11).

ACKNOWLEDGMENTS

I would like to thank A. C. Scott for many helpful discussions and for introducing me to this problem. Also, I would like to thank A. R. Bishop, T. Manteuffel, S. P. Layne, and R. R. Birge for stimulating discussions. The work was supported by the USDOE.

REFERENCES

1. See for example: "Physics in One Dimension," Eds. J. Bernasconi and T. Schneider, Springer Series in Solid-State Sciences, vol. 23, Springer-Verlag, New York (1981).
2. A. S. Davydov and N. I. Kislukha, Phys. Status Solidi (b), 59, 465 (1973) and Sov. Phys. JETP 44, 571 (1976).
3. A. S. Davydov, "Biology and Quantum Mechanics," Pergammon Press (1982).
4. T. Holstein, Mol. Cryst. Liq. Cryst. 77, 235 (1981).
5. G. Careri, U. Buontempo, F. Carta, E. Gratton, and A. C. Scott, Phys. Rev. Lett. (in press), see also A. C. Scott's contribution to these proceedings.
6. Yu. N. Chirgadze and N. A. Nevskaya, Biopolymers 15, 607 (1976).
7. W. H. Moore and S. Krimm, Biopolymers 15, 2439 (1976).
8. N. B. Abbott and A. Elliott, Proc. Roy. Soc. (London) A 234, 247 (1956).
9. J. M. Hyman in "Nonlinear problems: Present and Future," Eds. A. R. Bishop, D. K. Campbell, B. Nicolaenko, North Holland Publishing Company (1982) p. 91.
10. C. J. Brown and D. E. C. Carbridge, Acta Cryst. 7, 711 (1954). I thank L. MacNeil for pointing this reference out to me.
11. The coordinates used here were kindly supplied by F. Poulsen at the Carlsberg Lab., Copenhagen.
12. D. J. Thouless, Phys. Rep. 13, 93 (1973).
13. M. Karplus and J. A. McCammon, C. R. C. Crit. Rev. Biochem. 9, 293 (1981).

SOLITONS IN SYNTHETIC AND BIOLOGICAL POLYMERS

A. R. Bishop

Theoretical Division and Center for Nonlinear Studies
Los Alamos National Laboratory
Los Alamos, NM 87545

1. INTRODUCTION

Nonlinear science has, in recent years, begun to receive truly interdisciplinary attention and to involve a supporting interplay of analysis, computation and experiment. Most of the problems being addressed have a long history but are now benefiting from this new interdisciplinary view. The synergistic impact of computers plays an increasingly important role and there are some new concepts --solitons, toplogy, "universal" routes to chaos and its characterization, pattern selection and evolution, etc. Significant advances have occurred in our appreciation for the consequences of strongly nonlinear phenomena and our ability to experimentally detect them. In particular the "soliton paradigm"[1] has acquired, in the space of a decade, an astonishing list of applications across the natural sciences. Our focus here is only a small subset of these applications but already vast: namely applications in solid state materials and of those primarily low-dimensional examples (weakly coupled chains or layers). Furthermore, we will not discuss any problems arising in nonlinear diffusion equations, although these are fundamental in their own right for descriptions of reaction-diffusion systems, interface dynamics, nerve-pulse propagation, etc.[2]

Given the biophysical emphasis of this conference, our main concern is to indicate points of contact with the lessons which have been learned (sometimes painfully) in solid state and statistical physics over the last few years. To date only limited cross-fertilization has occurred between the nonlinear problems in the synthetic and biological literatures. Yet the parallels can be striking at all levels - mathematical equivalences and computational approaches and, even more importantly, phenomenological implications and suggested experimental probes. Since space is limited we can mostly only list some primary soliton concepts and material applications and give a guide to the relevant existing literature, whilst indicating potential applications in biophysical contexts (§§3-5). By way of illustration we will give a little more discussion of nonlinear effects at phase transitions (§3) and in certain synthetic polymers (§4). The relevant biophysic literature is very well described elsewhere in these proceedings, which we will therefore merely reference as appropriate.

2. SOLITON CONCEPTS AND APPLICATIONS

Good introductions to soliton mathematics include refs. 3-5. Solid state and statistical physics applications continue to expand rapidly but representative examples and surveys can be found in refs. 1,2, 6-8. With regard to potential applications in biophysics we would particularly emphasize the following general points:

(i) Solitons in one space dimension and with a small number of degrees of freedom at each "site" of a notional lattice, occur in only three forms -- kinks, pulses and envelopes.[3-5]

The first two are clear and typified by kink solutions to the sine Gordon (SG) and related equations, or pulse solutions to the Toda lattice or Korteweg-de Vries (KdV) equations. Both are solitons in that they are spatially localized but can propagate without change of shape. Indeed the pulse is just a derivative of the kink. Their physical origins can be very different, however, since their topologies are quite distinct -- the kink interpolates between two unconnected degenerate ground states whereas the pulse does not. Kinks are familiar as dislocations or domain walls, pulses as water wave solitons. Envelope solitons (often called "breathers" or "bions") are sometimes found a more difficult concept and yet in many ways they represent the most important and wide-spread class. Here the envelope prescribes the localized soliton structure which may be moving or not, but there is an additional periodic oscillation of the envelope amplitude or internal carrier wave. In these cases four (not two) canonical variables are needed to specify the soliton's state. Breathers are so important because they can interpolate between linear modes and extremely nonlinear structures such as kinks. Similarly they can always be reached by a suitably high order anharmonic expansion in the linear modes. Thus they lie naturally at the center of controversies in many fields disputing the advantages of conventional anharmonic perturbation theory versus particle-like soliton gas phenomenologies -- as with the wave-corpuscular nature of light it all depends on the property to be described (and the strength of the nonlinearity). The SG and nonlinear (cubic) Schrödinger (NLS) equations are typical sources of breathers.

(ii) Solitons in nature will rarely if ever be close to the precise objects conceived by mathematicians because of perturbation and/or intrinsic terms which destroy exact integrability of the governing equations of motion (as in the ϕ^4 equation used in §3.) Nevertheless, the basic balances which are present in true soliton equation can act to "label" generic types of soliton systems (and the solitons which follow) and also are often very robust against quite severe perturbations. Thus the generic ingredients of the SG

$$\phi_{xx} - \phi_{tt} = \sin \phi \tag{1}$$

and related equations (ϕ is a field variable; x and t are space and time) are wave-like propagation in the presence of a periodic local potential -- the latter is motivated from a "pinning", "locking", "registering" potential, depending on the physical origin[1,6-8] (dislocations in a metal, domain walls in a ferromagnet or ferroelectric, fluxons on a Josephson transmission line, discommensurations in a surface epitaxial layer, charged dislocations in a charge-density-wave material, etc.). The KdV equation

$$\phi_t + k\phi\phi_x + \phi_{xxx} = 0 \tag{2}$$

(k constant) characteristically combines weak nonlinearity and weak dispersion. In the case of the NLS equation

$$\psi_{xx} + i\psi_t + k|\psi|^2\psi = 0 \tag{3}$$

we always have in mind slow, long-wavelength self-modulation of an almost monochromatic wave with linear dispersion and weak nonlinearity. In eqn. (3), ψ is a complex (i.e. two-component) field and k is a positive constant for most cases of interest. (The negative sign generates quite different solitons and has fewer physical applications).

(iii) The stability against many perturbations (e.g. damping mechanisms, driving fields, impurites, grain boundaries, lattice discreteness, etc.) has naturally lead to the prevalent notion of solitons as "particles", i.e. collective modes, responding to perturbations primarily through a collective co-ordinate describing a center-of-mass motion.[1,3-8] While this view has much validity and leads to the importance of solitons for dynamics (e.g. as fluxons on Josephson transmission lines), the rigid, dynamic particle picture should not be used too literally. The equal importance of "solitons" for energetics, structure, statistical properties, etc., does not necessarily require simple dynamics. Thus the deformable nature of solitons (since they are not point particles) should always be considered. Again the effects of a discrete lattice are inherent in solid state and biological applications and act to impede or even freeze ("pin") soliton motion -- indeed this pinning is the origin of a low frequency resonance identified by SCOTT in the model of alpha-helix proteins discussed in these proceedings by LAYNE and LOMDAHL and SCOTT. There are many other examples in solid state -- the Peierls-Nabarro pinning barrier is well-documented in the theory of dislocations, for example.[9] More recently solitons appearing as "discommensurations" in incommensurate solid state phases can (on very long timescales) be randomly ("chaotically") trapped by a lattice pinning potential which overcomes long-range repulsions between discommensurations.[9]

It is particularly important to realize that transport can involve solitons fundamentally without implying simple ballistic or diffusive soliton motion, even though this is the most popularized mechanism --generalizing the familiar examples of slippage in metals via the motion of dislocations, or fluxon propagation on Josephson transmission lines, or soliton propagation along optical fibers. Hopping motion (of solitons or charges they may carry) can be of equally practical concern (as is likely in the conducting polymers of §4, where solitons appear as various kinds of "polarons", familiar in their own right). In other cases transport only occurs when solitons overlap sufficiently;[9] in such cases an independent soliton picture is quite irrelevant.

(iv) Some of the above remarks will have made it clear that nomenclature is unsettled.[1-8] The mathematician's soliton equations have very precise meaning related to their integrability.[3-5] This is directly connected with the remarkable soliton properties of collision stability, S-matrix factorability, Hamiltonian separability etc. One extremely striking consequence is the ability to quantize solitons exactly and construct quantum statistical mechanics.[4] It is now appreciated that this intimately connects soliton systems to those solvable by "Bethe Ansatz" techniques -- indeed it is clear that, via appropriate mappings, quantum soliton systems can be related to almost all exactly solvable problems in many-body and statistical physics and field theory (mostly one-dimensional quantum or two-dimensional classical systems). The importance of exactly solvable models (defining anchor points of precise knowledge about which a fabric of intuition and approximations can be woven) clearly gives true soliton systems a major unifying importance. However the term "soliton" has come to be used much less precisely in most physics literature -- signifying any spatially localized, finite energy, dynamic or static, intrinsic order-parameter configuration (often termed an inhomogeneous state). Stability is usually guaranteed on topological grounds; non-topological solitons (e.g. the pulses or envelopes above) are stabilized by dynamics or external influences (e.g. impurities, boundaries).

Hopefully nomenclature will evolve to label a polaron a polaron or a vortex a vortex, etc. but the "soliton" label does serve a useful purpose to focus on an important nonliner paradigm. It is this paradigmatic sense for solitons which is most relevant in both solid state and biophysics.

(v) Applications of paradigmatic (and in a few cases nearly literal) solitons in solid state all center on transport, energetic, or structural properties. The solitons may be excitations (thermal, quantum, critical fluctuations) or ground state structures. In some cases they may be much less populous than linear modes and yet, for appropriate physical properties (e.g. scattering functions, transport) much more important. Earlier references will indicate the immense scope for the soliton paradigm, but the most direct and striking solid state examples have probably occurred in low-dimensions. Examples include[6-8] Josephson transmission lines, certain magnetic chain materials, mercury chain compounds, fast-ion conductors, structurally distortive compounds, etc. There are many other potential applications but the most intensive investigations currently emphasize electronically and magnetically active polymers and charge-transfer salts.[10] Examples include spin- and charge-density-wave materials (KCP, TCNQ salts, $(TMTSF)_2PF_6$, $NbSe_2$, $NbSe_3$, ...), piezo- and pyro-electric polymers (PVF_2), and conducting polymers (§4).

It is natural to look also for applications in biology, prompted both by the frequently low-dimensional (polymer) settings and by basic chemical and structural similarites with low-dimensional solid state systems. For instance, the hydrogen bond plays a pervasive role in biophysics which leads us to expect similarities with structurally phase-transforming materials such as ferroelectrics (there are even some quasi-1-dimensional examples such as CsD_2PO_4, acetanilide, nylon-66, certain polydiacetylenes). Many of the questions posed in DNA are essentially ones of (linear and nonlinear) lattice dynamics (albeit in complicated lattice structures) which have been faced recently in a number of structurally transforming materials (see §3). Ionic and protonic transport mechanics as well as charge-transfer processes are increasingly discussed in biophysics literature but are extremely similar to processes in fast-ion conductors or charge-transfer compounds. Peptide ($H \cdots N\text{-}C = 0$) units occurring in ordered (e.g. α-helix proteins or synthetic polypeptides chains) or disordered (e.g. globular proteins) environments have been discussed extensively in terms of self-trapping of lattice vibrational energy ("excitons") because of nonlinear exciton-phonon coupling. As we shall see this self-trapping produces "solitons" of great similarity to self-localized electrons ("polarons") in synthetic metals (§4), where *electron*-phonon coupling operates. We list further nonlinear similarites between synthetic-and bio-polymers in §5.

3. NONLINEAR LATTICE DYNAMICS AND STRUCTURAL PHASE TRANSITIONS

Strong nonlinearity is important in all types of phase transitions[1,6-8] -- from traditional first order (where "droplets" are relevant) and continuous (below) transitions to more recently considered commensurate-incommensurate (below), 2-dimensional melting and martensitic classes. We briefly discuss one class which should have biophysical relevance:

Order-Disorder and Displacive Transitions.

Observing structural transformations in ferro- (and anti-ferro-) distortive (as well as ferroelectric) crystals has a long history, but the role of intrinsic nonlinearity has only recently begun to be confirmed experimentally or theoretically.[11] Key experimental observations may be summarized as *precursor* effects (mixed or heterophase fluctuations; the occurence of *clusters* of the "wrong" phase for $T \gtrless T_c$, the critical temperature), *central peaks*, and *soft modes*. Emphasizing order-parameter structures (such as short-range clusters) represents a complementary concern to that of "critical exponents" and can be a considerable conceptual tool when visualizing critical or sub-critical dynamics and constructing theoretical frameworks. The temperature regime where nonlinear (cluster) effects are most pronounced is greatest in *lower* dimensions.

The simplest model Hamiltonian we might consider has the familiar double-well on-site potential (Ginzburg-Landau expansion) form.[11] In one dimension (1-d) we write this as:

$$H[U] = \sum_{i=1}^{N} \left[\frac{1}{2} m \left(\frac{\partial U_i}{\partial t} \right)^2 + \frac{1}{2} A U_i^2 + \frac{1}{4} B U_i^4 + C(U_{i+1} - U_i)^2 \right] , \qquad (4)$$

where m is particle mass and the order parameter U_i might be, e.g., rotation, displacement (for example of a light mobile lattice relative to an immobile heavy ion or fictitious reference lattice). For simplicity we take $A<o$, $B>o$, $C>o$ and all constant. We then have a crude model of a uniaxial unstable lattice system with harmonic ferrodistortive coupling and local double-well poentials at each lattice site (Fig. 1a). At high temperatures we can expect that the particles will oscillate above the double-well

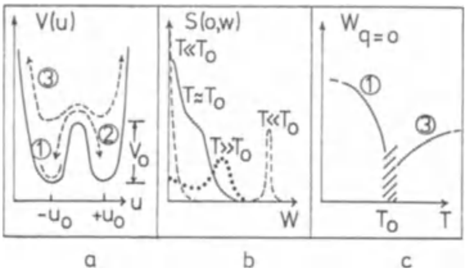


Fig. 1 (Schematic). (a) Local ϕ^4 potential in Hamiltonian (4). Stable modes are 1) low-T "phonons", 2) domain walls, 3) high -T phonons; (b) soft mode-central peak characteristics for ϕ^4 model; (c) softening of q = 0 phonon (incomplete in 1 - d). Hatched region may sustain 1) and 3). Type 2) responses appear for $T \lesssim T_0$, yielding a central peak.

structure with approximately uniform probability distribution but at low temperatures they will preferentially occupy one of the degenerate wells. Notice that we do not include a mean-field "Ginzburg-Landau" temperature dependence in A of the form $A = a(T-T_o)$. We have in mind situations where T_o is a high mean field scale temperature and we operate at $T << T_o$. The model will then have a true transition temperature T_c quite independent of T_o (in fact $T_c = o$ in 1-d). There is no difficulty including thermodynamic field dependences in the coefficients and this is sometimes essential -- as in mean field descriptions of multiphase equilibria or discommensuration defects.[9,11] Model (4) is of course highly oversimplified in several respects. It is, however, able to isolate dominantly nonlinear effects and considerable generalizations are possible, particularly within molecular dynamics simulations.

If variations in U are small on the scale of a lattice spacing $\ell \equiv x_{i+1}-x_i$, then a continuum representation is valid: $U_i-U_{i+1} \cong \ell^2(\partial U/\partial x)^2$. This situation will hold when the intersite coupling is strong, $C > |A|$, and we term it the "displacive" limit. Kink widths will be $>>\ell$ and the kink enerby E_K will be an "activation energy" for $k_BT \leq E_K$. The opposite regime, $C << A$, can be termed "order-disorder"; kinks will be sharp and particles distributed with a thermal randomization and activation energy $|V|$ (Fig. 1a). As we approach this limit, the discrete lattice can severely modify the excitation dynamics. In the displacive regime we can introduce a continuum Hamiltonian density and the coupled Euler-Lagrange equations of motion for $\{U_i\}$ reduce to a partial differential equation for U(x). From (4)[1] we find

$$m\, U_{tt} - m\, C_o^2 U_{xx} - |A|U + BU^3 = 0, \tag{5}$$

where $C_o^2 = 2C\ell^2/m$ is the harmonic sound velocity for this system. We refer to (4) and (5) as the "ϕ^4" Hamiltonian and nonlinear equation respectively.

Traditionally, a sharp distinction has been drawn between order-disorder and displacive systems. However, it is now appreciated that these are merely extremes of a general class. Indeed sufficiently close to T_c there is a crossover to universal order-disorder behavior in all cases -- as is confirmed[6,8,11] by analytic approaches and supported by molecular dynamics simulations as well as real experiments.

Two characteristic temperatures have to be appreciated. The first is the true long-range ordering temperature -- $T_c = 0$ in strict 1-d and small for anisotropic coupled chains. A second temperature, T_o, is not a true critical temperature but marks an incipient first order "condensation" or "local ordering" -- for $T>T_o$ particles have sufficient energy to oscillate above the on-site double well structure, whereas for $T < T_o$ they prefer to oscillate in an individual displaced well; however for $T > T_c$ $\langle U\rangle = 0$ so "clusters" of particles oscillating in the same well are separated from neighboring clusters in the other well by "domian walls" (our kinks in the 1-d problem). The temperature T_o can be characterized in variou equivalent ways. The self-consistent phonon approximation (SCPA) finds the optimal harmonic description for particle oscillations and at a temperature $\simeq T_o$ predicts a first order transition from in-well to above-well oscillations. (i.e. $T_o \leftrightarrow$ "soft mode temperature"). It is incapable of accounting for the kink solutions which are the reason for an incipient condensation only. Alternatively, one can study the single-particle probability distribution and identify T_o with the temperature below which a double hump distribution develops. More sensitive many-particle distributions must be studied[11] for isotropic higher-d. We find according to these criterion: $k_BT_o \sim 2|V_o|(C/|A|)^{\frac{1}{2}}$. A very reasonable interpretation of T_o is that it is the temperature at which the kink excitations become thermally unstable. To operate this criterion we need to include a "thermally renormalized" kink energy at high T. This can be accomplished approximately by comparing with static correlation function estimates, which suggest[12]

$$k_BT_o \simeq E_K(T_o) \simeq 0.4\ E_K(0) \simeq 2|V_o|(C/|A|)^{\frac{1}{2}} \qquad (6)$$

The "scaling" dependence on $(A/C)^{\frac{1}{2}}$ follows automatically from the displacive approximation (5).

The SCPA has enjoyed popularity as a soft mode description for displacive systems. It can indeed be successful in a substantial range of (T, ω, q) (temperature, frequency, wave-vector) space but fails in certain regimes. The SCPA replaces $-\frac{1}{2}|A|U_i^2 + \frac{1}{4}BU_i^4$ (equ. (4)) with $\frac{1}{2}(-|A| + \frac{1}{2}B\langle U^2\rangle)\ U_i^2 = \frac{1}{2}A^*(T-T_o)U_i^2$. Thus distinct sectors of excitations space (below) are omitted. Calculations of the dynamic structure factor S(q,ω) in 1-d can be interpreted to show that (i) the SCPA is valid for all T if $q \gtrsim \ell^{-1}(|A|/C)^{\frac{1}{2}} \propto d^{-1}$ (the kink width 2d increases as C/A increases); (ii) it is valid for $T \gtrsim T_o$; the SCPA dispersion,

$$\omega_q^2 = \omega_{sm}^2 + (2C/M)\ (1-\cos q\ell),\ \text{with}\ \omega_{sm}^2 = (|A|/M)(3\langle U^2\rangle/\langle U^2\rangle-1)$$

is only valid for $\omega_{sm} > \omega^*$ where ω^* is smaller for more displacive systems (C/A increasing). Remark (iii) is the fact that there is only an incipient soft mode. In fact the strongly anharmonic regions where SCPA fails for 1-d are characterized by the dominance of kink excitations and the appearance of a central peak associated with them.

In the displacive 1-d case in particular, a quite complete phenomenology can be constructed from a knowledge of the general traveling wave solutions to the continuum equation (3), i.e. available lattice dynamic modes. The general solutions appear, see e.g. Ref. 13, in the form of Jacobi elliptic functions. The linearly stable types are illustrated in Fig. (1a). Physically, they represent (i) "low-T" or "in-well" phonons; (ii) "high-T" or "above-well" phonons ("phonon" here includes general anharmonic periodic solutions); (iii) domain-walls or kinks. These excitations give a better description of the transition region $T \simeq T_o$ than the enforced harmonic modes of the SCPA. Snapshot distributions are not transparently "separable" for $T \gtrsim T_o$ but response functions are much more so. As AUBRY has emphasized,[13] the mean energy per unit length ε_o, for the three excitations as a function of ω and q, is an important quantity. See Fig. 1 of Ref. 13.

Note the softening of $\omega(q=0)$ for low- and high-T phonons as $\varepsilon_o \to |V_o|$ (Fig. 1a). It remains to relate temperature to ε_o. In a non-integrable anharmonic system equipartition of energy between modes has to be an approximate procedure. However it is known that for $q \gtrsim q_o$ (above) the available motions are essentially harmonic at all T, so that one can conjecture that the energy per particle to be associated with large amplitude anharmonic motions, $E_a(T)$, is

$$E_a(T) \simeq k_B T \left[1 - q_{bz}^{-1} \int_{q_o}^{q_{bz}} dq\right] \propto (A/C)^{\frac{1}{2}} k_B T \quad .$$

This procedure works remarkably well in explaining the location of strong response characteristics at all T: a mode will be thermally active (with signature in $S(q,\omega)$) if $\varepsilon_o(\omega,q) \lesssim E_a(T)$. Also note that $E_a(T) \simeq |V_o|$, which from Fig. 1 of Ref. 13 we expect to be associated with the soft mode-central peak onset, implies $T \simeq T_o \simeq E_K(T_o)$ as suggested in (6). To include broadening of resonances in $S(q,\omega)$ we need in general to improve this independent excitation gas phenomenology by including interactions.

Dynamics correlation functions (and the Fourier transform $S(q,\omega)$) cannot be calculated analytically in the strongly anharmonic region, however the combination of phenomenology (above) and molecular dynamics leads unambiguously[11-13] to the picture shown schematically in Fig. 1. At low T most motions are low amplitude, and response characteristics of an extended periodic mode ("low-T phonon") appear. This softens as $T \to T_o$ and is replaced at higher T with the "high-T phonon" response. The kink excitations give rise to the very different central peak response. This appears as T is lowered as a broad, low intensity resonance around the incipient soft mode temperature T_o and becomes sharper and more intense as T is further lowered, diverging at T = 0 (the 1-d critical temperature). The central peak appears strongly at q = 0 and disappears for $q \gtrsim q_o$. It is solely due to the kink excitations in the simpler models, and should be thought of as a dynamic and imcomplete (i.e. having width in ω and q space) Bragg peak. The central peak width is partly determined by the inverse average kink separation and effective kink velocity: for $T \simeq T_o$ there are numerous stable, mobile kinks but none at T = 0.

In isotropic 2- or 3-d, $|T_o - T_c|$ is substantially smaller than in 1-d and therefore it has been difficult to observe cluster properties experimentally: an excellent history is given by MÜLLER.[8] Success has come recently using EPR measurements which discriminate between the fast phonon ($\sim 10^{-11}$ sec) and slow cluster ($\sim 10^{-7}$ sec) timescales. Real anisotropic materials such as polymers offer special advantages of a substantial T_c and large $|T_c - T_o|$.[14] It is important to emphasize that clusters are most important physically because of their impact on dynamics and timescales (as with "critical slowing down"). Central peaks, clusters, and slow time-scales can have other origins,[11] for example static or dynamic impurities which may trap intrinsic clusters or induce similar distortions -- clearly these intrinsic and extrinsic mechanisms are not mutually exclusive but part of a single unified picture.

There are now many fine examples of structurally transforming materials including hydrogen-bonded ferroelectrics (e.g. KH_2PO_4).[11] Quasi-one-dimensional examples include KCP, TTF-TCNQ, TaS_3, $NbSe_3$, CsH_2PO_4 (see refs. 10).

In biological contexts one can anticipate numerous examples of ferroelectric, distortive and melting phase transitions -- some examples have been discussed elsewhere.[15] One of the most topical discussions centers on transitions between structures in DNA, e.g. the A and B phases.[15] It is not our intention to describe the existing literature. This is already reviewed elsewhere in these proceedings. However, we do emphasize that the lessons learned in the conventional solid state contexts will surely be paralleled in biopolymers, especially because of their effective low-dimensionality. We must of course recognize that the "crystals" involved are extremely complex, and information (theoretically or experimentally) on lattice structure, lattice dynamics, mode softening, etc., is still at a relatively primitive stage.[15] Nevertheless, we can expect that a proper appreciation of the complementary roles of intrinsic local, large-amplitude conformational distortions and quasi-harmonic extended modes and extrinsic defects will be achieved much more rapidly if the history of soft modes and central peaks is appreciated in the biophysics community.

Finally, we briefly mention one other general class of phase transitions in solid state contexts where strong nonlinearity has been extremely important, namely commensurate-incommensurate transitions.[9] These occur in many situations where there are competing interactions and periodicities and the ground state can be intrinsically inhomogeneous. Specifically, the system may homogeneously accommodate one interaction ("be commensurate") for some extended regions of space but bridge these regions with incommensurate defects ("discommensurations"), often described by solitons of the SG variety, eqn. (1). The density of these defects tends to zero as a transition is approached from the incommensurate side by varying temperature, pressure, etc. This situation had achieved little attention in solid state physics until the last few years, but is in fact very common -- the competing interactions might be an underlying lattice spacing, spin- or charge- or mass-density-wave period, etc., and the discommensurations might appear in the mass density, spin density, charge density, helical spin period, etc. Theoretical and experimental tools have sharpened rapidly and there are now many excellent observations of commensurate-incommensurate transitions and even of the discommensurations themselves.[9] Pinning of the discommensurations by the discreteness of the lattice or by extrinsic defects can even lead to random ("chaotic") ordering of the discommensurations and to many nearly equivalent metastable states.[9] Metastability and hysteresis are typical of this class of phenomena. It is very easy to imagine competitions between interactions in biophysical contexts which should fall into this general class and lead to intrinsic structural inhomogeneities. Indeed primitive arguments along these lines have already been put forward.[16]

4. SOLITONS IN POLYACETYLENE AND RELATED MATERIALS

Many organic and organo-metallic polymeric materials represent fine settings for the soliton paradigm with structural, energetic and transport consequences.[10] A particular example, which has received intensive theoretical and experimental attention is polyacetylene, $(CH)_x$, synthesized for example as a film which can be doped to near-metallic conductivity levels.[10]

We emphasize that the research effort devoted to $(CH)_x$ is driven primarily by its fascinating technological potential. Nevertheless it is also fair to claim that soliton concepts have modified the conceptual basis within which experiments and theories are designed in the strongly interdisciplinary field of "conducting polymers", which includes $(CH)_x$. In this brief report, we wish to develop the theoretical status of $(CH)_x$ modeling, because it illustrates many general features of soliton research such as the equivalences which solitons reveal between disparate physical contexts: here we will make use of connections with model field theories which are found[17] to have the same kink- and polaron-soliton states as predicted in $(CH)_x$. Likewise the exact solubility of these field theories (using soliton techniques) leads naturally to explicit statements about the structure and excitations of a wide range of other polymer models. Soluble models are of course useful here, as elsewhere, because they can bring physical clarity to complex features which are essentially preserved (e.g. because of symmetry) when exact solubility is not possible.

Descriptions of $(CH)_x$ from a chemical or physical perspective can be found elsewhere.[10] Briefly, the major synthesis effort is now devoted to controlled synthesis-morphology-property relation studies, both of $(CH)_x$ and many other members of the growing family of conducting polymers. It is unlikely that $(CH)_x$ will survive as the example of prime technological interest, but its extreme simplicity for modeling purposes has meant that it has played a primary stimulant role for exciting experiments and theories. Even for $(CH)_x$, new syntheses have emerged ranging from almost totally amorphous to near single crystal forms.[10]

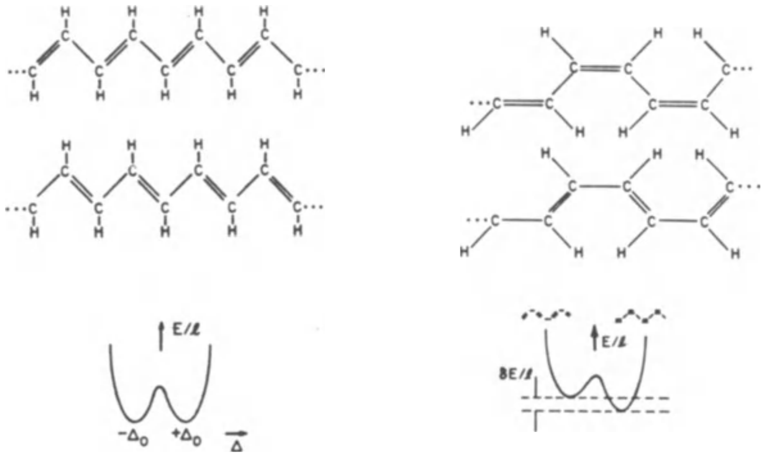


Fig. 2. Schematic bond structures for (a) *trans*-(CH)$_x$ and (b) *cis*-(CH)$_x$. Also included are schematic plots of energy per unit length vs. band gap parameter Δ for *uniform* Δ. Note the degeneracy in (a) corresponding to the two equivalent ground state conformations. In (b) the lack of degeneracy, δE/ℓ, has important consequences (§3).

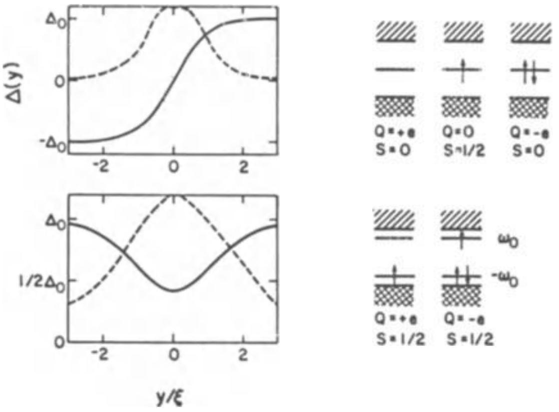


Fig. 3. Intrinsic defect states in *trans*-(CH)$_x$ and associated electronic levels for (a) a kink, (b) a polaron. Dashed lines indicate localized state probability densities. Q and S denote charge and spin, respectively, and the length scale $\xi_0 = V_F/\Delta_0$. Analytic formulae for the profiles and continuum state phase shifts are given in ref. 17. From the latter we find that a kink removes 1 state from both the conduction and valence bands, whereas the polaron removes 2. This immediately explains the Q-S relations.

For our present purposes, we wish merely to consider[17,18] an idealized $(CH)_x$ single chain and introduce some primitive chemical language for the two isomers of $(CH)_x$, which we term *cis* and *trans* for simplicity. The situation is summarized in Fig. 2. A uniform chain of (CH) monomers would have one unpaired π-electron per monomer (σ orbital bands lie far from the Fermi level and are not important here) and the polymer should be a conductor. The uniform chain is however unstable toward *dimerization* into the *trans*-(most stable) or *cis*-isomeric forms in which all electrons are saturated giving an insulator. Precisely this situation is described in solid state as a Peierls-Fröhlich distortion,[18] which for a half-filled π-band is a dimerization.

A crucial difference between *trans*- and *cis*-$(CH)_x$ is the presence of ground state *degeneracy* in the former, as depicted in Fig. 2a. This will lead to the possibility of "free" *kink-soliton* states in *trans*-$(CH)_x$ about which much has been written.[10,17,18] In contrast only bound kink-antikink (*polaron*)-solitons[17] are possible if there is no such degeneracy. The idea of conformational defects should be clear as distortions of bond lengths between possible ground state conformations -- either the same ground state or different ones (as for *trans*-$(CH)_x$. Imagine rupturing a double bond and separating the unpaired spins, leaving free radicals or, upon charging, carbanions or carbonium ions. Indeed the concept of S-½ conjugation defects in polymers is not at all new. The main addition from the soliton viewpoint is that the conformational defects are typically *extended* over many C-C lengths. This has important consequences for energetics and dynamics. The conformational defects are of course accompanied by defects in the electronic density as indicated below.

SU, SCHRIEFFER and HEEGER (SSH)[18] introduced a simple tight-binding Hamiltonian for a pure, isolated *trans*-$(CH)_x$ chain:

$$H = - \sum_{n,s} [t_o + \alpha(U_n - U_{n+1})][C^+_{ns} C_{n+1,s} + \mathrm{h.c.}] + \frac{1}{2} K \sum_n (U_n - U_{n+1})^2 + \frac{1}{2} M \sum_n \dot{U}_n^2 \quad . \tag{7}$$

Here C^+_{ns} creates an (π) electron of spin s at site n, U_n is the displacement of the (CH) unit at the nth site from its undistorted position, and M is the mass of a (CH) unit. Although very primitive this model has some very interesting properties because of the coupling (α) between the electronic and phonon degrees of freedom. There are of course additional effects neglected in (7) which can be quite important, e.g. interchain coupling or electron correlations. These are now the subjects of intensive studies[10] but we will not consider them here.

For a half-filled π-conduction band, SSH demonstrated in the *adiabatic* limit that (7) is spontaneously unstable towards one of two degenerate ground states, corresponding to the Peierls-dimerized A and B configurations in Fig. 2a -- distinctions need to be made between even and odd length chains and between sizes 4N and 4N+2 (N integer). In addition, they concluded numerically that excitations or, for electron concentrations close to half-filling, intrinsic defect states in the dimerized pattern appear as kink- or polaron-like conformational distortions, with important spin-charge relations and consequencies for doping or photo-generation of carriers. The ground state and excitation structure follow naturally in a *continuum* theory of (7) which is extremely accurate in the case of $(CH)_x$.

The continuum limit of (7) has been given by several authors and in a mean-field adiabatic approximation results in the following equations for the *static* 1-particle electron wave-fns.[17]

$$\begin{aligned} \varepsilon_n u_{ns}(y) &= -iv_F \frac{\partial}{\partial y} u_{ns}(y) + \Delta(y) v_{ns}(y) \\ \varepsilon_n v_{ns}(y) &= +iv_F \frac{\partial}{\partial y} v_{ns}(y) + \Delta(y) u_{ns}(y) \end{aligned} \tag{8}$$

and the self-consistent gap equation

$$\Delta(y) = \frac{-g^2}{\omega_Q^2} \sum_{ns}{}' \, [v^*_{ns}(y) u_{ns}(y) + u^*_{ns}(y) v_{ns}(y)] \quad . \tag{9}$$

Here y is the continuous variable, $\omega_Q^2/2g^2=K/4\alpha^2$, $v_F=2\ell t_0$ (ℓ is the undistorted lattice constant) $g=4\alpha(\ell/M)^{1/2}$, u and v are the two components of the electron spinor field linearized around the Fermi surface, and $\Delta(y)$ is proportional to the staggered lattice displacement $\tilde{U}_n=(-1)^n U_n$: $\Delta(y)=4\alpha\tilde{U}(y)$. The prime in (9) indicates summation over occupied states.

The problem posed by (8) and (9) is closely connected with "soliton" problems in many other areas, e.g. superconductivity and nonlinear optics.[17] Alternatively, a simple transformation casts (8) into the form of single particle Dirac equations for massless fermions in a potential $\Delta(y)$, and this problem can be shown[17] to be equivalent to one example of a soluble field theory of interacting massless fermions. *Soliton*-like properties permit the construction of analytic, closed-form expressions for *all static* configurations -- the dimerized ground state, a kink (K) or a polaron (kink-antikink, $K\bar{K}$, bound state). These are illustrated in Fig. 3. The kink and polaron should be viewed as localized conformational defects with associated localized electronic levels. Note the unusual spin-charge relations for K or $\bar{K}$, Fig. 3.

There a number of generalizations of the soluble model outlined above. One important family of soluble models was introduced by BRAZOVSKII.[19] The ordered A=B alloy and the SSH model with broken ground state degeneracy (as in cis-$(CH)_x$) represent two popular members of the family. The 50:50 alloy is intriguing because of the possibility of a mechanism for fractionally charged kink-solitons.[20] Here Hamiltonian (7) is modified by an alternating atomic orbital on-site energy. Materials such as (poly)carbonitrile are candidates for this model.

The SSH model with groundstate symmetry-breaking (c.f. Fig. 2b) is of extremely general importance: other Peierls-distorted systems with multiple ground states are certainly possible but they are far outnumbered by materials with a unique ground state plus one or more metastable conformations. Some of these have potential as practical conducting polymers (e.g. polypyrroles, poly(para) phenylenes, polydiacetylenes, polythiophenes). The central point is clear and independent of specific modeling. Namely, the energy difference between the unique ground state and any metastable conformations will provide a linear "confinement" potential - imagine trying to create kink and anti-kink and to separate them. This means that kinks cannot be "free" (on the pure chain) and must bind in pairs, i.e. as polarons. The bi-polarons (i.e. two self-trapped charges) will be wider than polarons but not unstable as in trans $-(CH)_x$. This general idea has now been appreciated in modeling of the more complicated polymers above.[21] Note that bipolarons can have the same signatures as charged kinks, i.e. charged but spin-0. An explicitly soluble model (using the same soliton techniques as earlier) can be constructed[17,19] which demonstrates all the confinement features. Introducing $\Delta(y)=\Delta(y)+\Delta_e$, where Δ_e is a constant symmetry-breaking term, results in analytic polaron solutions which have precisely the same functional form as for the trans model (7) but with location of the gap states (c.f. Fig. 3) at $\pm\omega_o$ with $\omega_o/\tilde{\Delta}_o=\cos\theta$ and $\gamma\tan\theta=(\pi/4)(n^+-n^-+2)\theta$, with $\gamma=\Delta_e/\lambda\tilde{\Delta}_o$, $n^\pm$ the occupations of the gap states, and $\lambda^{-1}=\pi v_F\omega_Q^2/2g^2$ (a dimensionless coupling constant).[17]

The Hamiltonian (7) describes a situation of inter-molecular electron-phonon coupling. In many organic and organo-metallic materials, intra-molecular modes are much more numerous and can be at least as important. In fact inter- and intra-molecular mode couplings may well be in competition and this has lead to suggestions[22] of interesting phase diagrams where they are operative together -- both in terms of the allowed ground states and excitations. The simplest model of electrons coupled to intra-molecular phonons is perhaps the "molecular crystal model", familiar in solid state for many years,[23] which in 1-d takes the form[24]

$$
\begin{aligned}
H = {} & \frac{1}{2}M \sum_n \dot{y}_n^2 + \frac{1}{2}M\omega_o^2 \sum_n y_n^2 \\
& - J \sum_{n,s} (a^+_{n,s} a_{n+1,s} + a^+_{n+1,s} a_{n,s}) \qquad (10) \\
& - A \sum_{n,s} y_{n,s} a^+_{n,s} a_{n,s}
\end{aligned}
$$

In (10), y_n is a normal coordinate usually referring to some internal vibration of a molecular unit at the n-th site, with an associated mass M and natural frequency ω_0. $a^+_{n,s}$ ($a_{n,s}$) creates (annihilates) an electron with spin s at the n-th site. The four terms in (10) represent, respectively, the lattice kinetic energy, the vibrational energy of the molecular lattice, the electronic kinetic energy associated with moving an electron between (nearest neighbor) sites, and the coupling of the electron and lattice (phonon) motions.

It will be clear how the terms inter- and intra- arise from a comparison of (7) and (10). We have contrasted these models in detail elsewhere.[24] Here we note only that (i) the molecular crystal model was developed[23] explicitly as a generic model for "self-localized (or trapped)" electrons, i.e. polarons, and (ii) despite the explicit differences between (7) and (10), the polarons which they can support are quite similar. In fact, in the weakly-bound polaron limit (i.e. low-amplitude, spatially extended), the two polarons become indentical and are described (in a continuum theory by the NLS equation (3).[24] Even more interestingly we note that the same tight-binding model (10) can be, and has been, applied to the self-localization not only of electrons but also magnons and vibrational quanta (sometimes called "excitons"). In this case a^+_n creates a magnetic or vibrational quanta and the constant A refers to the strength of magnon-phonon or exciton-phonon coupling. Note that there is a difference of statistics between electrons (fermions) and magnons or excitons (bosons).

We stress this last point because precisely the same model Hamiltonian has recently been proposed in a biological context without recognizing the very relevant solid state polaron literature -- and therefore the possibility of short-circuiting analysis, the recognition of pitfalls, and the choice of appropriate experiments. We refer to the subject of self-localized excitations ("solitons" of the NLS variety) due to exciton-phonon interactions in polypeptide chains, i.e. H•••N - C = O coupled units. The coupling of the longitudinal phonons to peptide vibrations has been suggested both in an inter[25] and intra[26] form. The biological contexts may also have lessons for the synthetic ones. For example polypeptide chains occur as 3 coupled strands in α-helix proteins[25] (myosins, e-coli, mitochondria, etc.) and the biological discussions have also focussed attention on synthetic polymers with 2-coupled strands (e.g. acetanilide)[26] and nearly isolated strands (e.g. in nylon-66 and certain polydiacetylene side groups). The *coupling* between chains has revealed a variety of chain-*sharing* localized exciations. The question of interchain coupling is only now receiving corresponding attention in materials such as conducting polymers (above). Again the combined presence of inter- and intra-molecular phonon coupling has so far been discussed for *static* excitations in the solid state literature,[22] whereas the more interesting question of *dynamics* has been motivated by the biological concerns -- leading, for example, to the suggestion[26] of intra-molecular coupling acting as a transient (picosecond scale) self-trapping mechanism nucleating activity in the intermolecular modes on a longer time scale. This should have wider applications. On the other hand, the effect of *disorder* (of the lattice locations and/or site energies in (10)) has already been considered in solid state situations with electron- (or magnon- or exciton-) phonon coupling -- e.g. in amorphous semiconductors. From these studies it is already clear that disorder *alone* can lead to self-localization of the elementary excitations as well as to extended states. (This is the famous theory of "Anderson localization".)[27] Since the coupling to phonons can also lead to self-trapping[23,28] (less readily in $d > 1$ than $d = 1$), these effects tend to reinforce each other, and there has been a protracted "chicken-and-egg" discussion about their relative importance.[28] This history will necessarily be repeated in the descriptions of coexisting localized and extended stationary states in globular (disordered) proteins, such as the lysozyme discussed by LOMDAHL. The functional roles for self-trapped excitations in biology remain to be clarified but they lend themselves to fascinating speculations.[15,25]

It should be clear that some of the most intriguing functional consequences (e.g. energy or charge transduction) for localized ("soliton") objects concern their influence on transport. This is true just as much for conducting synthetic polymers as for biopolymers. However, the story is very far from complete in either case. We reemphasize that simple diffusive soliton transport will probably have limited relevance in either case. In general we can expect (thermal or quantum) tunneling to play a major role -- either of the total soliton entity or of the self-trapped excitation -- leading to "hopping" transport mechanisms.[23,27,28] In addition, recent numerical studies[29] of semi-classical soliton dynamics have revealed unexpected subtleties, even for so simple a model as the SSH, eqn. (7). For example, the solitons have a *maximum* velocity which is unrelated to the sound speed. Again, "breather"solitons (strongly anharmonic phonons), analagous to the dynamic solitons of the NLS equation (3), are readily excited (e.g. by laser stimulation). Soliton dynamics and transport in polymers leave, much to be revealed!

5. SUMMARY

There has been little space in this article to mention the many connections between solid state (e.g. synthetic polymer) and biophysics, with strong nonlinearity playing the unifying role. Striking parallels occur at all important levels: analytical (as in the case of self-trapping mechanisms, §4); numerical techniques; experimental implications, applications, and techniques. (At once, conventional solid state probes are at last being devoted to biopolymers but their complexity is demanding new extremes and techniques -- in a real sense, biopolymers are a new frontier of materials science).

As important as the above connections is the commonality of phenomena. These include: hydrogen-bonding (compare peptide chains or DNA or cellulose with hydrogen-bonded ferroelectrics);[15] structural (and ferro-electric) phase transitions (3§); competing interactions (leading, for example, to commensurate-incommensurate phase transitions and intrinsic structural "chaos", §3); thermal or quantum nucleation (which may be relevant to premelting or intercalation centers in DNA (see SOBELL and ref. 15); self-localization phenomena (§4); transport mechanisms (diffusion or hopping of self-localized states, proton transport (e.g. through biomembranes), ionic diffusion or hopping, charge-transfer, etc., compared with polaron transport in molecular crystals and conducting polymers or transport in fast-ion conductors); piezoelectricity (familiar in synthetic polymers such as poly(vinylidene) fluoride and suggested in materials such as DNA); macromolecular architecture (e.g. in lipid bilayers) and its synthetic parallel in solid state polymerizations such as in single crystal polydiacetylenes.[30]

There is no space here to give due attention to all these topics. However we can conclude by reiterating two dominant themes: (i) solitons are important in condensed matter, especially in reduced dimensionality, for structure, energetics and transport; and (ii) an interdisciplinary approach to strong nonlinearity has had profound success in many other areas of the natural sciences. It will surely yield similar benefits in biophysics, perhaps with even more significant consequences.

References

1. A. R. Bishop, J. A. Krumhansl and S. E. Trullinger, Physica D 1, 1 (1980).

2. "Nonlinear Problems" Present and Future," Eds. A. R. Bishop, D. K. Campbell and B. Nicolaenko, North-Holland, Mathematics Studies 61 (1982).

3. G. L. Lamb, Jr., "Elements of Soliton Theory," Wiley Interscience (1980).

4. "Solitons," Eds. R. K. Bullough and P. J. Caudrey, Springer-Verlag (1980).

5. R. K. Dodd, et. al., "Solitons and Nonlinear Wave Equations," Academic Press (1983).

6. "Solitons and Condensed Matter Physics," Eds. A. R. Bishop and T. Schneider, Springer-Verlag, Solid-State Sciences 8 (1981).

7. "Physics in One Dimension," Eds. J. Bernasconi and T. Schneider, Springer-Verlag, Solid-State Sciences 23 (1981).

8. "Statics and Dynamics of Nonlinear Systems," Eds. G. Benedek, H. Bilz and R. Zeyher, Springer-Verlag, Solid-State Sciences 47 (1983).

9. e. g. S. Aubry, in "Order in Chaos," Physica D 7 (1983); P. Bak, Rep. Prog. Phys. 45, 587 (1982).

10. e. g. Journal de Physique, Colloque C3, Supplément au no. 6, Tome 44 (1983).

11. A. D. Bruce and R. A. Cowley, "Structural Phase Transitions," Taylor and Francis, Monographs on Physics (1981).

12. T. R. Koehler, A. R. Bishop, J. A. Krumhansl and J. R. Schrieffer, Solid State Commun. 17, 1515 (1975); A. R. Bishop, in "Proceedings of the International Conference on Lattice Dynamics," Paris 1977, Ed. M. Balkanski, Flammarian Press (Paris) (1978).

13. S. A. Aubry, J. Chem. Phys. 64, 3392 (1976).

14. A. R. Bishop and J. A. Krumhansl, Phys. Rev. B 12, 2824 (1975); A. R. Bishop and W. C. Kerr, preprint (1983).

15. e.g. B. F. Putnam, E. W. Prohofsky and L. L. van Zandt, Biopolymers 21, 885 (1982); H. Bilz, H. Büttner and H. Fröhlich, Z. Naturforsch 36b, 208 (1981); J. A. Krumhansl and D. M. Alexander, "Structure and Dynamics: Nucleic Acids and Proteins", Eds. E. Clementi and R. H. Sarma, Adenine Press (1983), p. 61.

16. e.g. H. Toyoki, et al., preprint (1983).

17. D.K. Campbell and A.R. Bishop, Nucl. Phys. B200, 297 (1982); see also refs. 2, 8.

18. W.P. Su, J.R. Schrieffer and A.J. Heeger, Phys. Rev. B22, 2099 (1980).

19. S.A. Brazovskii and N.N. Kirova, JETP Lett. 33, 4 (1981); see also ref. 10.

20. M.J. Rice and E.J. Mele, Phys. Rev. Lett. 49, 1455 (1982).

21. J.C. Scott et al., preprint (1983); T.C. Chung, et al. preprint (1983).

22. S. Kivelson, Phys. Rev. B 28, 2653 (1983); W.P. Su, preprint (1983)

23. T.D. Holstein, Annal. Phys. 8, 325 (1959); D. Emin and T.D. Holstein, ibid. 53, 439 (1969); T.D. Holstein, Mol. Cryst. Liq. Cryst. 77, 235 (1981).

24. D.K. Campbell, A.R. Bishop and K. Fesser, Phys. Rev. B 26, 6873 (1982).

25. A. S. Davydov and N.I. Kislukha, Sov. Phys. JETP 44, 571 (1976); A.S. Davydov, "Biology and Quantum Mechanics), Pergamon Press (1982); A.C. Scott, Phys. Rev. A 26, 578 (1982).

26. G. Careri, et al., Phys. Rev. Lett. 51, 304 (1983); see also A.C. Scott, these proceedings.

27. see D.J. Thonless, Phys. Rep. 13, 93 (1973); N.F. Mott and E.A. Davis, "Electronic Processes in Non-crystalline Materials," Oxford University Press (1971).

28. e.g. P.W. Anderson, Nature (Physical Sciences) 235, 163 (1972); D. Emin, Physics Today, June (1982), p. 34.

29. A.R. Bishop, et al., Phys. Rev. Lett. 52 (in press) and Synmetals (in press).

30. e.g. G. Wegner, in "Molecular Metals," Ed. W.A. Hatfield, Plenum Press (1981).

PICOSECOND RELAXATIONS IN PROTEINS AND BIOPOLYMERS OBSERVED BY MM-WAVE SPECTROSCOPY

F. Kremer, A. Poglitsch and L. Genzel

Max-Planck-Institut für Festkörperforschung
Heisenbergstraße 1, 7000 Stuttgart 80
Fed.Rep.Germany

ABSTRACT

Broadband absorption measurements at millimeter-wave frequencies (40 GHz-170 GHz) are reported for several biological materials (lyophylized haemoglobin, lysozyme, silk keratin, poly-L-alanine). The measurements were extended over the temperature range from 4.2 K to 300 K using the novel oversized-cavity-technique (OCT). The millimeter-wave absorption of the dried materials was found to increase nearly exponentially with temperature and approximately as $\nu^{1.5}$ with frequency. This frequency and temperature dependence is quantitatively described as being due to three distinct relaxation processes on a picosecond time scale occuring in asymmetric double well potentials. These processes are most probably assigned to the NH...OC hydrogen bonds of the peptide backbone. For lysozyme the influence of hydration was investigated. The contribution of adsorbed water can be distinguished from the intrinsic proccesses by its different frequency dependence.

INTRODUCTION

Dynamic processes (e.g. relaxations, librations, vibrations) in biological macromolecules involving collective motions of individual side chains or more extended parts of the molecule play an important role in their biological functioning and, for enzymes, in the enzyme-substrate interaction.[1-3] In contrast to X-ray investigations which detect displacements averaged in time, dielectric measurements can give direct information about the typical times being involved in dynamic processes. At mm-wave frequencies ($3\text{x}10^{10}$-$3\text{x}10^{11}$ Hz), which correspond in time to picosecond processes, relaxations and vibrations in biological macromolecules are predicted.[4,5]

EXPERIMENTAL

Various experimental reasons are responsible for the lack of information especially in the mm-region. There is firstly the low absorption of biomolecules in contrast to the high absorption of liquid water which prevents at present the possibility of obtaining useful results from aqueous solutions. Secondly there is the difficulty of handling solid samples in the small-sized single mode waveguide or resonator-systems which are conventionally used in the mm-region. Furthermore, temperature variation over a wide range causes problems due to the different thermal expansion coefficients of the samples and the wave-guide material. Finally quasi-optical systems using mirrors, lenses and horn antennae suffer from disturbing standing waves and frequency-dependent diffraction effects when coherent microwave sources are used. We have, therefore, employed the rather novel technique of the untuned, oversized high-Q cavity[6-10] for conducting the mm-wave spectroscopy. This method measures only the absorption of a sample inserted into a cryostat inside the cavity, but is essentially not disturbed by scattering and diffraction. The fact that the untuned cavity is a broadband device enables us to carry out measurements in the frequency region from 40 to 150 GHz. A specially designed fused-silica cryostat allowed the temperature range from 4 K to 300 K to be covered.[8] Further measurements at 10 GHz were made with the usual cavity perturbation technique.

In order to determine the absorption coefficient α of the material the oversized cavity technique requires an assumed value for n, the real part of the complex index of refraction. For all samples under investigation a value of n = 1.6 was used. It could be shown that a variation of n within limits of 0.2 will change the calculated α-value by less than 2% only.

RESULTS

Dried Materials: The respective absorption coefficients α of the dried materials (haemoglobin, lysozyme, silk keratin, poly-L-alanine) were found to exhibit similar temperature dependences (Fig.1 and 2), namely a nearly exponential increase with temperature from 50 K to 300 K. Below 50 K, α drops quickly and seems to level off at about 10 K. In the range of the exponential increase with temperature, α is about linear in frequency, while at low frequencies α is proportional to ν^2. We did not observe any sharp features in the spectrum which could be discussed in terms of vibrational modes.

Hydrated Lysozyme: The influence of hydration at a fixed frequency of 50 GHz is shown in Fig.5. At temperatures below 120 K the absorption is only slightly affected by bound water, while at higher temperatures a strong increase is observed that is proportional to the water content. The difference in the dielectric function between the dried and the hydrated samples turned out to be nearly frequency independent.

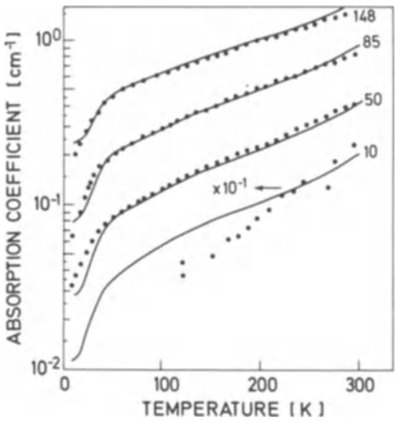


Fig. 1: Absorption coefficient vs. temperature of dried haemoglobin for the frequencies of 10 Ghz, 50 GHz, 85 GHz and 148 GHz. The dots are the experimental values. The solid lines are fitted curves according to eq. 7. The measurement at 10 GHz was carried out using the cavity-perturbation-technique. The measurements at mm-wave-frequencies were made using the oversized cavity-technique.

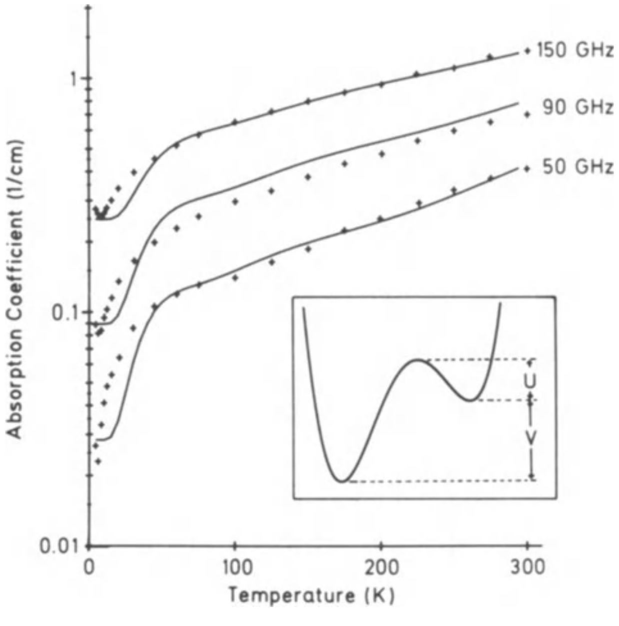


Fig. 2: Absorption coefficient vs, temperature for dried lysozyme (water content < 0.5 g(H_2O)/g(protein).[11] Inset: Asymmetric double well potential as used for the relaxation model. Otherwise as Fig. 1.

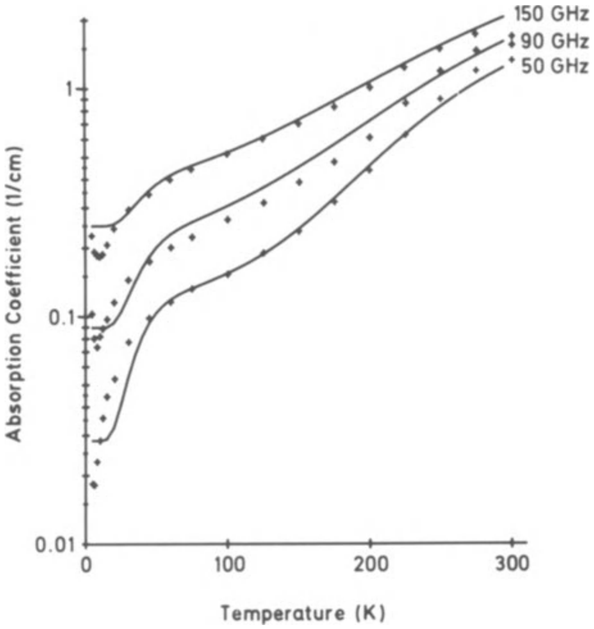


Fig.3: Absorption coefficient vs. temperature for hydrated lysozyme (18% g(H_2O)/g (protein). The contribution of bound water, mainly occuring above 150 K is nearly frequency independent as found by comparison with the dried sample.[11] Otherwise as Fig.1.

DISCUSSION

In subsequent discussion we show that a minimal set of three relaxation processes, occuring in asymmetric double well potentials, fits the data. The observed frequency and temperature dependence of the absorption coefficient cannot be explained by mechanisms such as vibrational excitations from a thermally excited ground state, two-level resonant absorption as observed in amorphous materials, multi-phonon absorption or a Debye relaxation process of the type:

$$\varepsilon(\nu,T) = \varepsilon_\infty + (\varepsilon_0 - \varepsilon_\infty)(1 - i\nu\nu_\tau^{-1})^{-1}. \qquad (1)$$

Here ε_∞ is the high-frequency dielectric constant (DC) and ε_0 the DC below the frequency regime of the relaxation process in question. ν_τ is the relaxation frequency which is usually assumed to have the following temperature-dependence of solids:[12-14]

$$\nu_\tau = \nu_\infty \exp(-U/kT). \qquad (2)$$

U is a potential barrier which has to be overcome thermally by the system in order to change from one stable configuration to another of the same energy implying, therefore, a symmetrical double-well potential model. The term $(\varepsilon_0 - \varepsilon_\infty)$ in eq.(1) is approximately given by the Kirkwood-Fröhlich relation:[11-14]

$$\varepsilon_0 - \varepsilon_\infty = 4\pi(kT)^{-1}\ \varepsilon_s(\varepsilon_\infty + 2\varepsilon_s)^{-1}\ ((\varepsilon_\infty + 2)/3)^2 N\mu^2 \equiv CT^{-1}. \quad (3)$$

Where ε_s is the static DC, N the spatial density of the double-well systems and μ the absolute value of the vectorial difference between the dipole moments in the two wells. The absorption coefficient follows from eq.(1):

$$\alpha = 2\pi\nu\varepsilon_2(nc)^{-1} = (\varepsilon_0 - \varepsilon_\infty)(nc)^{-1} 2\pi\nu^2\nu_\tau(\nu^2 + \nu_\tau^{\ 2})^{-1}\ . \qquad (4)$$

In this equation n is the refractive index, c the speed of light and ε_2 the imaginary part of $\hat{\varepsilon}$. Eq.(4) with eq.(3) can be discussed for two limiting cases: a) $\nu^2 \ll \nu_\tau^{\ 2}$, in this case one finds α to be proportional to $\nu^2\nu_\tau^{-1}T^{-1}$ yielding with eq.(2) a strong decrease of the absorption coefficient with T and an increase with ν; b) $\nu^2 \gg \nu_\tau^2$, in this case α is found to be protortional to $\nu_\tau T^{-1}$ and hence no longer frequency dependent. Evidently, both cases are in contradiction with the experimental findings but assuming the relaxation processes occur in asymmetric double well potentials (s.inset Fig.2) one finds[15] instead of eq.(3):

$$\varepsilon_0 - \varepsilon_\infty = CT^{-1}\ \exp(-V/kT)(1 + \exp(-V/kT))^{-2} \qquad (5)$$

and instead of eq.(2)

$$\nu_\tau = (\nu_\infty/2)\ \exp(-U/kT)(1 + \exp(-V/kT). \qquad (6)$$

Inserting eq.(5) and (6) into eq.(4), one can get the desired increase of α with T even in the case $\nu^2 < \nu_\tau^2$, retaining thus the frequency-dependence if $V > U$. It turns out, however, that a single process of this kind is still not able to describe the observed $\alpha(\nu,T)$ over the hole range of variables in question. In such cases a continuous distribution of relaxation frequencies or potentials V and U is often considered. In order to avoid the arbitrariness of fitting functions we used instead a minimal set of three relaxation terms which were able to fit the data.

The observed leveling off of the $\alpha(T)$ curves at low T together with the observed ν^2-dependence can be explained by a temperature-independent wing of a strong and broad vibrational band found for proteins as well as for poly-L-alanine in the FIR between 120 and 150 cm^{-1}.[16] With this, our dielectric function for fitting the data of haemoglobin can be written:

$$\varepsilon(\nu,T) = \varepsilon_\infty + \frac{S_o\nu_o^2}{\nu_o^2-\nu^2-i\nu\gamma_o} + \sum_{j=1}^{3}\frac{C_j}{T}\,\frac{e^{-V_j/kT}}{(1+e^{-V_j/kT})^2}\,\frac{1}{1-i\nu/\nu_{\tau j}} \qquad (7)$$

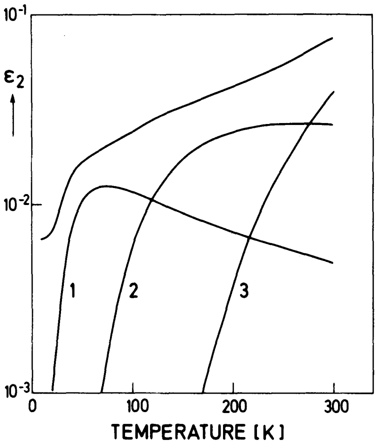


Fig.4 The temperature dependence of the imaginary part ε_2 of haemoglobin at 85 GHz. Curves 1-3 correspond to the relaxation processes 1-3 respectively. The total dielectric function is composed of the three relaxation processes plus a vibrational term which causes the leveling off at low temperatures.[15]

As an example of the final fitting for haemoglobin the three relaxation processes which, when summed, form the resultant $\varepsilon_2(T)$ curve are shown in Fig.4. In Tab.1 the fitting parameters are compiled. The full lines in Fig.1 demonstrate how well the experimental $\alpha(T,\nu)$ data is represented.

Table 1: Parameters for fitting the mm-wave data of haemoglobin
It is assumed that $N = 6x10^{21} cm^{-3}$.

Relaxation	ν_∞ (GHz)	U (kcal/Mol)	V (kcal/Mol)	$N\mu^2$ (cgs)	μ (Debye)
1	370	0.04	0.25	$2.8.10^{-16}$	0.21 ± 0.01
2	350	0.20	0.91	$2.1.10^{-15}$	0.59 ± 0.006
3	300	0.40	3.2	$8.1.10^{-14}$	3.7 ± 0.4

FIR-vibration: $\varepsilon_\infty = 1.77$, $\nu_0 = 4400$ GHz $\simeq 147$ cm^{-1}
$S_0 = 0.385$ $\gamma_0 = 3879$ GHz

Differential-phonon-processes could also cause the observed frequency and temperature dependence to a certain extent.[17] They are characterized by a frequency dependence which is at least quadratic. This is in contrast to the experimental finding of a frequency dependence proportional to about $\nu^{1.2}$, but it is possible that this mechanism contributes slightly to the observed absorption.

The influence of hydration at a fixed frequency of 50 GHz is shown in Fig.5. At temperatures below 120 K, the absorption is only slightly affected by bound water, while at higher temperatures a strong increase is observed. The frequency dependence of this absorption process is obtained from Fig.3 showing the absorption coefficient of hydrated lysozyme (18% H_2O) for three frequencies. It turns out, that the difference in the dielectric function between the dried sample and the hydrated one is nearly frequency independent. For a relaxation process the absorption coefficient, as given by

$$\alpha = \frac{\omega\varepsilon_2}{nc} = \frac{\Delta\varepsilon}{nc}\,\frac{\omega^2\tau}{1+\omega^2\tau^2}\,, \tag{8}$$

is frequency independent in the case where

$$\omega^2\tau^2 >> 1 \text{ or } \nu >> 1/2\pi\tau. \tag{9}$$

Within the accuracy of the measurement, the relaxation rate $1/2\pi\tau$ cannot exceed 10 GHz at room temperature, corresponding to a relaxation time $\tau < 1.6x10^{-11}$s and hence no dispersion at mm-wave frequencies. This allows one to distinguish the absorption of bound water clearly from the fast intrinsic processes of the anhydrous protein and to describe it by a further relaxation process, which is proportional in strength to the water content, occuring in a symmetric double well potential with a potential barrier of 2.2 kcal/mol (1100 K).

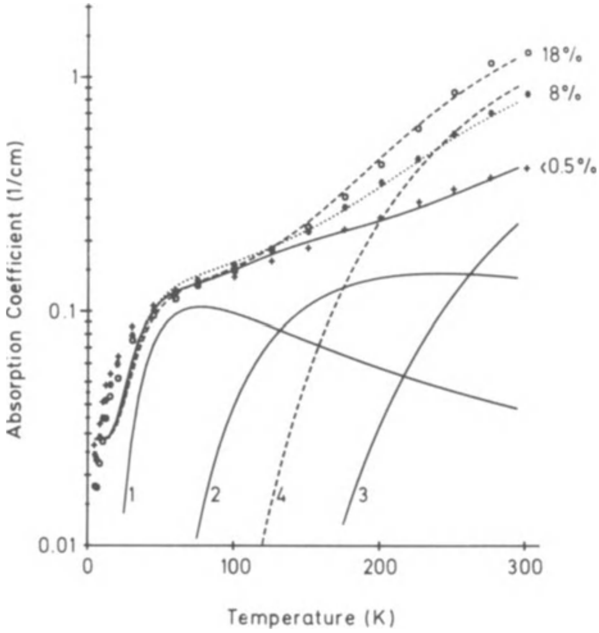


Fig.5 Absorption coefficient vs. temperature of lysozyme at 50 GHz for three hydration levels.[11] The dielectric function of the dried material is composed of three relaxation processes (solid lines 1-3). The contribution of bound water is represented by a further relaxation, which is shown for 18% hydration as the dashed line 4.

To yield the observed polarization for the third relaxation process, the dipole moment of a bound water molecule in the dried sample (water content > 0.5% w/w) would have to be $5\mathrm{x}10^{-29}$Cm (15 Debye), which is a totally unrealistic value.[18] So we can conclude, that the intrinsic relaxations 1-3 cannot be due to bound water. A relaxation of polar side groups, as the dominant mechanisms in this frequency range, can be excluded too, because poly-L-alanin, which contains no polar side groups, exhibits a quite similar frequency and temperature dependence of the absorption coefficient at the same order of magnitude.[15]

What is the microscopic nature of the assumed relaxation processes 1-3? (See table 1 and Fig.4 for haemoglobin; Fig.5 for lysozyme). Comparing the values of ν_∞, the high temperature relaxation frequency (see eq.6), one finds them rather similar. The resulting relaxation times are in the order of picoseconds. This indicates relaxations of small molecular subunits. The depth H = U+V of the potential well for the third relaxation process is 3.6 kcal/mol (1800 K), a value which is slightly below the binding energies of hydrogen bonds. Therefore, relaxation process 3 is considered to be a short-term disruption of H-bonds. From the known molecular weight of haemoglobin and its density the number of amino-acid residues per molecule (each of which has one NH...OC bridge) was found to be $N = 6\mathrm{x}10^{21}$ cm^{-3}. From that the dipole moment change μ involved in the relaxation process can be calculated as $\mu = 3.7$ Debye. It is interesting to note that the NH...OC bridge itself has a static dipole moment of 3.7 Debye. The fact that the disrupted hydrogen bond has a certain dipole moment as well might point towards some cooperativity between neighbouring bridges.

The discussion of the nature of the relaxations 1 and 2 must be more speculative. An indication that they might occur also on the NH...OC bridges is given by the fact that these bridges do not form a straight line. The oxygen atom with its local charge of about 0.38 e forms a dipole moment of the order of 1 Debye to the line connecting N and C. One could, therefore, consider the relaxations 1 and 2 as being caused by asymmetric double-well potentials lying perpendicular to the H-bridge. It is worth noting that the observed relaxation effects, found here in solid and dry samples, are expected to exist also for the materials in aqueous solution as long as the submolecular units (H-bridges) in question are in the hydrophobic interior of the biopolymer and are thus not exposed to the outer water. For lysozyme and poly-L-alanine similar values for the relaxation frequency, the potentials and the changes of the dipole moment are found. This similarity is a strong indication that the observed processes are located within the peptide backbone. The observed short relaxation times, if relaxation is the dominant absorption-mechanism, make the assignment to a process involving the proton within the NH...OC group highly probable.

REFERENCES

1. P.J. Artymiuk, C.C.F. Blake, D.E.P. Grace, S.J. Oatley, D.C. Phillips, and M.J.E. Sternberg (1979). Nature 280, 563
2. H. Fraunfelder, G.A. Petsko, and D. Tsernoglou (1979). Nature 280, 558
3. R. Huber and W.S. Bennett (1983). Biopolymers 22, 261
4. J.A. McCammon, B.R. Gelin, and M. Karplus (1976). Nature 262, 325

5. M. Karplus, J.A. McCammon (1979). Nature 277, 578
6. D.T. Llewellyn-Jones, R.J. Knight, P.H. Moffat, H.A. Gebbie (1980), 535
7. F. Kremer, J.R. Izatt, Int.J.of Infrared and Millimeter Waves, (1981), 2, 675
8. F. Kremer, L. Genzel, "Application of Untuned Cavities for Millimeter Wave Spectroscopy", Proc.6th Int.Conf.on Infrared and Millimeter Waves, Miami, Dec.1981, paper T-1-4.
9. J.R. Izatt, F. Kremer, (1981) Appl.Optics 20, 2555
10. F. Kremer, (1983) "Biological Effects and Dosimetry of Non-ionizing Radiation" Ed.M.Grandolfo, S.M. Michaelson, H.Rindi, Plenum, 233
11. A. Poglitsch, F. Kremer, L. Genzel (1983). submitted to J.Mol.Biol.
12. H. Fröhlich, Theory of Dielectrics (1958), Clarendon Press
13. C.J.F. Böttcher, P. Bordewijk, Theory of Electric Polarization (1973), Elsevier, Amsterdam
14. R. Pethig, Dielectric and Electric Properties of Biological Materials (1979), John Wiley and Sons, Chichester
15. L. Genzel, F. Kremer, A. Poglitsch, G. Bechtold (1983), Biopolymers, 22, 1715
16. L. Genzel, F. Kremer, A. Poglitsch, G. Bechtold (1983), in "Coherent Excitations in Biological Systems" Ed.H.Fröhlich, F. Kremer, Springer Verlag, Berlin Heidelberg, New York, Tokyo, 58
17. R. Stolen, K. Dransfeld (1965), Phys.Rev. 139, No.4A, 1295
18. S. Bone and R. Pethig (1982). J.Mol.Biol. 157.

NONLINEAR MECHANISMS IN MOLECULAR ENERGY TRANSFER

T.W. Barrett

Surface Chemistry Branch, Code 6170, Chemistry Division
Naval Research Laboratory, Washington, DC 20375

The anharmonic oscillator theory is applied to the nonuniformly doped, electrically conducting, trans-polyacetylene $(CH)_x$ chain, as well as the long Josephson junction with varying bias current. The variations in bias current determine the number of fluxons on the Josephson junction. As the soliton is the equivalent of the fluxon in trans-$(CH)_x$, the average voltage on the $(CH)_x$ chain is demonstrated to be a function of propagation velocity and the number of solitons contained on it.

The number of solitons on any segment of trans $(CH)_x$ is a function of the monomer energy gap which is determined by doping. A nonuniformly doped polyacetylene chain made up of such segments is shown to exhibit an average voltage $\langle v \rangle = \langle n \rangle u/\ell$ where u is a constant propagation velocity, ℓ is the length of the polymer chain and $\langle n \rangle$ is the average number of solitons on the chain. The experimental quadratic current-voltage relations reported for trans-$(CH)_x$ support this conclusion.

I. PRELIMINARIES

The notion of a propagating soliton is not a simple notion and the requirements for its existence are rigorous. The aim of this paper is to indicate the complexity involved in its existence and to suggest criteria for such existence. another aim is to indicate the fundamental importance of the concept of soliton, first called such in 1965 (1, 2), but whose origins stem from observations by John Scott Russell (3). The timeliness of this endeavor is due to the consideration being given on the one hand, to proteins (4-10) and doxyribonucleic acid, DNA (11), as mediators of vibrational-type solitons, and, on the other hand, to trans-polyacetylene, $(CH)_x$, as a mediator of soliton dependent electrical conductivity or charge-domain-wall movement (12, 13). In this paper, the charge-domain-wall soliton is used as a working model for the physics of the vibrational-type solitons postulated in proteins and DNA.

The charge-domain-wall-type soliton of $(CH)_x$ has received much attention. The case for the existence of such solitons in $(CH)_x$. has been made mainly on the basis of electron spin resonance (ESR) data (14). That solitons do, in fact, exist in trans-polyacetylene can also be treated as a hypothesis. On the other hand, the mathematical entity of the soliton has undergone considerable evolution and its history has been described by Scott (15). The intention here, is not to question whether the mathematical entity of soliton propagation can exist in polymers (undoubtedly it does), but to refine judgment concerning the necessary conditions for its existence. This is a matter of discrimination based on both physical and mathematical reasoning.

All physical situations considered here are described in terms of a classical anharmonic oscillator moving in an anharmonic double potential well (DPW). As the soliton is known to exist in the line of pendula joined with elastic coupling, and also in the long Josephson junction as the fluxon, the attitude has been taken here that if, indeed, the soliton exists in polymers, whether of the vibrational or charge-domain-wall-type, then it should conform to certain established criteria of behavior already rigorously studied in the aforementioned situations.

INTRODUCTION

Trans-polyacetylene, $(CH)_x$ is electrically conducting and the underlying mechanism mediating this conduction is thought to be due to the presence of dynamic force relations describable by the mathematical entity of the soliton (14). There are a number of current-voltage relations reported for $(CH)_x$ which show a quadratic dependence of current on the voltage (16-23) (cf. Figure 1). With the dopant concentrations $y > 2x10^{-3}$ the EPR signal is significantly reduced from that of the undoped (and electrically nonconducting) trans-$(CH)_x$, and for $y > 2x10^{-2}$ the signal vanishes (24-25), indicating that the dopant binds to free electrons providing the conditions for propagation of charged excitations (soliton) which do not transport electrons. This doping is, however, not expected to be uniform throughout any $(CH)_x$ chain.

There have been a considerable number of theoretical treatments of soliton propagation in $(CH)_x$ (12-13, 26-27). A different point of view is considered here, permitted by the NMR observations of Nechtschein et al. (28) on changes in the proton relaxation rate upon doping of $(CH)_x$ and by Chen et al. (29) that doping adds to the double bond and at the highest concentration results in the formation of a substituted polyethylene. These observations suggest that the doping process results in a competition between the dopant and the hydrogen (state of resonance) for a carbon bond.

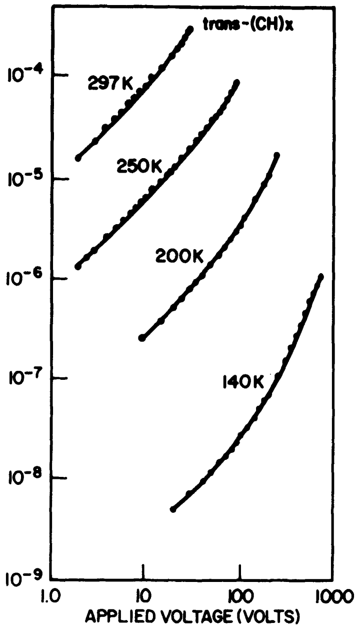


Figure 1. Current-voltage characteristics of trans-$(CH)_x$ at various temperatures. From ref. (23).

As electrical conductivity in $(CH)_x$ is dependent upon this doping, considered nonuniform, the possible reasons are examined here for the quadratic current-voltage relations in trans-$(CH)_x$ using a soliton charge-carrying mechanism based on the anharmonic oscillator model (30).

The model presented is microscopic in that the current-voltage relation and resistance effects are treated as properties of the carbon bond-charge of the dopant interaction. There are also macroscopic effects due to the fact that $(CH)_x$ chains are finite and charged species must get from one chain to the next. The model presented below, does not take into account these macroscopic effects, which undoubtedly influence the overall magnitude of electrical resistance in a "rope" of intertwined $(CH)_x$ chains, but not the quadratic dependence of the I-V relation, which is of central interest in this paper.

For purposes of exposition $(CH)_x$ chains are shown in which *every* carbon atom has a dopant molecule nearby with which it interacts as well as interacting with a hydrogen molecule. However, in practice, current will flow in $(CH)_x$ for one dopant every 100 carbon atoms. The examples, shown below, have 100% doping merely for purposes of exposition.

There is also a simplification in the use of a double potential well for the competitive interaction of the hydrogen and dopant molecules for the carbon atom. Such potentials are more complex, but the real complexity has no bearing upon the qualitative result obtained.

ANHARMONIC OSCILLATOR MODEL

The anharmonic oscillator model is based on an equation of motion for a monomer of the general form (31,32):

$$\dot{\phi} = x^3 + \frac{Ax}{B} + C = 0, \tag{1A}$$

where x is the B position of the carbon double bond and ϕ is a configurational potential (Figure 2), and for coupling to 2d immediate neighbors:

$E_W = |A|^2/4B$ is the depth of the local potential, (1B)

$E_B = C|\overline{A}|/Bd$ is the representative energy of the (1C)

carbon bond. EB defines the isomerized state which can take on one of two positions.

$h = E_W/E_B$ is a ratio and C is defined in Equ. (2) below. (1D)

For h>>1 only localized transitions occur, but for h<<1 a displacive transition (32-39) results in an incommensurate phase domain wall diffusion (soliton propagation).

To describe a propagating transition in a chain of monomers of the type depicted in Figure 2A, the transition becomes displacive (40-43) (h<<1), resulting in domain wall diffusion. The potential energy of the Hamiltonian description of the process is (31,32):

$$V \doteq \frac{C}{2} \frac{n.n.}{\ell.\ell'} (u(\ell) - u(\ell')^2) + \sum_{\ell} \dot{\phi}(x(\ell)), \tag{2}$$

where u(ℓ) describes the displacement of the bond defining the isomerized state along the length of the polymer chain, ℓ , n.n. indicates nearest neighbor and ϕ is defined by Equ. (1A). Thus, whereas the monomer description, Equ. (1), resembles the Su, Schrieffer and Heeger (12,13) description, the displacive chain description, Equ. (2), does not, as Su et al. adopted a many-body approach which precluded a description of the displacement transition inherent in soliton propagation.

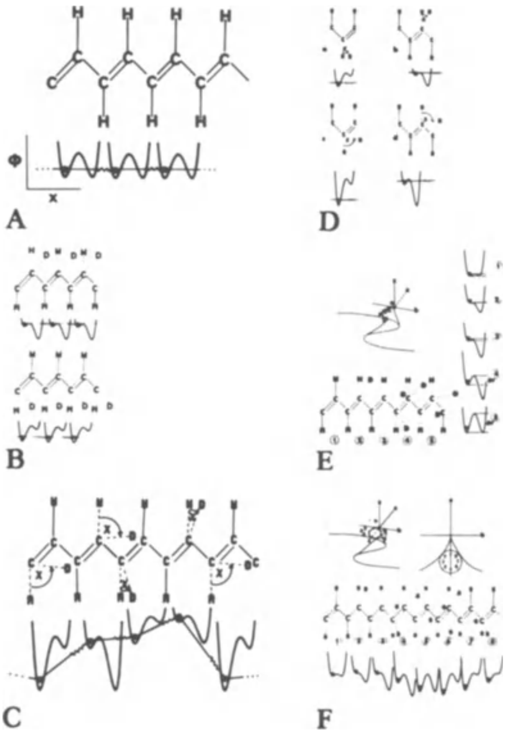


Figure 2. A. Potential energy (Φ) versus double bond position (x) diagram for (CH)$_x$ chain. B. (CH)$_x$ chain, with two different cases of doping. The dopants (D) represents the effective sum charge of single or many dopant species acting on the carbon atoms of the (CH)$_x$ chain. C. and D. Inhomogeneous doping. E. and F. For explanation see text.

Taking into consideration that the doping process results in the previously mentioned competition with the hydrogen (state of resonance) for a carbon bond (29), then ideally the doping results in two types of configuration, A and B (Figure 2B). In these figures the dopants (D) represent the effective sum charge of possibly single or many dopant species acting on the carbon atoms of the $(CH)_x$ chain. They thus represent the possibly multiple static charge influences on the carbon bond transfer involved in changing from one isomerized state to the other. In the Figure 2B (top) doping condition, the bond transfer from left to right (effecting a transfer from one isomerized state to the other) is facilitated, as the dopant provides a stronger electron acceptance in the isomerized state to be adopted. In the Figure 2B (bottom) doping condition, on the other hand, the bond transfer from left to right is less likely, as the dopant provides a stronger electron acceptance in the isomerized state already existing. The actual result of doping is more probably the nonideal forms shown in Figure 2C where the angle χ is a pictorial representation for the degree of resonance interaction of bond 1 with bond 2 and the depth of the configurational double potential well (Figure 2D).

It is a consequence of the anharmonic oscillator model that any of the joined monomer segments shown in Figure 2C may be represented as a trajectory over the a,b surface representation of the monomer element phase space described by Equ. (1). In Figure 2E is shown an arbitrary semi-circular trajectory; and Figure 2F depicts the possible nonuniform doping of $(CH)_x$ and monomer configurational potential wells for a complete circular trajectory. The particular trajectory is, of course, arbitrary, but for purposes of exposition the circular trajectory will be used here. In the case of trajectory 1 → 8, at each of the 8 monomer sites and for an increasing applied current to the $(CH)_x$ chain, the measured voltage, if it could be measured, would show an initial increase to V and then greatly diminishing increases after that value. Conversely, for an increasing applied voltage to the $(CH)_x$ chain, the measured current, if it could be measured, would show an initial slow increase until a ΔI value is reached and then increase greatly.

With constant current applied and the resistance and voltage measured at each monomer location, if such could be measured, the relation shown in Figure 3A is obtained. For the particular nonuniformly doped chain considered, the monomer current-voltage relations at each monomer site is shown in Figure 3B.

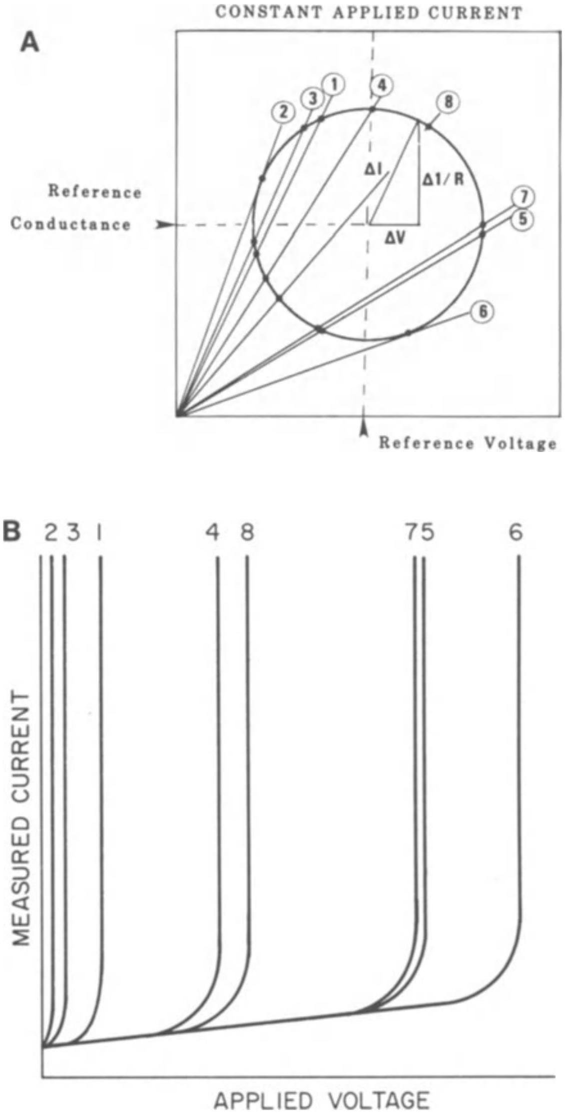


Figure 3. A. Logarithmic representation of ΔV, 1/ΔR and constant applied ΔI relations of the trajectory of Figure 2F. The numbers refer to the locations in Figures 2E and 2F. B. Current singularities in the nonuniformly doped $(CH)^x$ chain for the locations in Figures 2E and 2F.

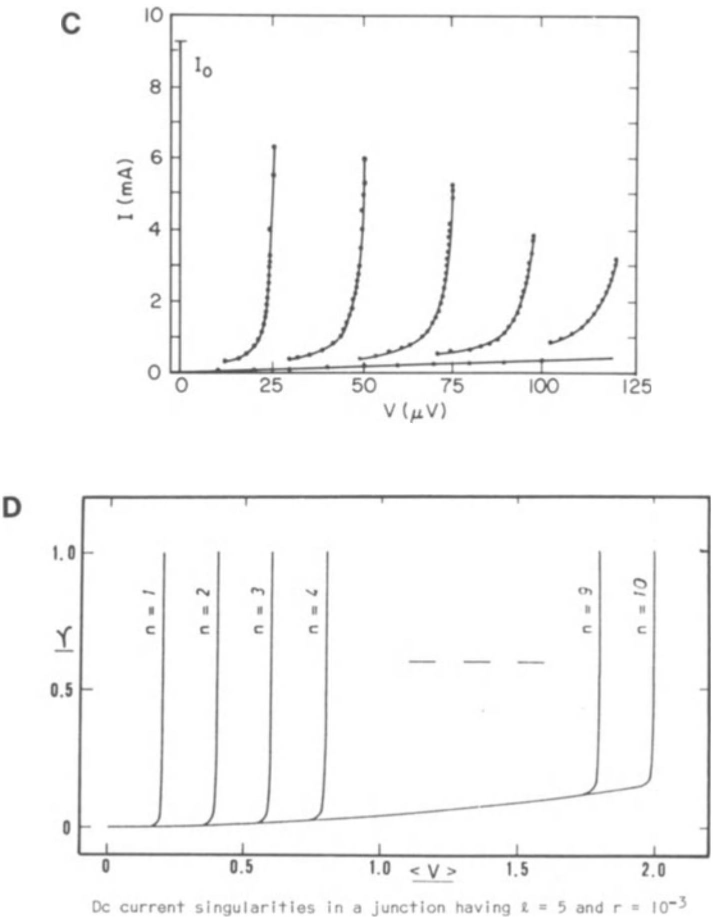


Figure 3. C. Current-voltage characteristic of a very long Josephson junction showing zero-field current steps in the presence of an external magnetic field applied transversely across the junction. From ref. (48). D. Current-voltage characteristics for n fluxon oscillations. From ref. (45). Fluxon oscillations move to one end of a long Josephson junction are reflected as an antifluxon and move to the other end, etc. (52). Such oscillations can absorb power from an external source of bias current. Costabile et al. (47) have also made similar calculations using an approximation scheme rather than an exact solution with periodic boundary conditions.

There is little hope of measuring such monomer current-voltage relations presently for technical reasons. However, the task to relate the soliton-mediated conductivity at each monomer site to the conductivity whicy may be measured for a collection of such chains may be approached analytically. For example, $(CH)_x$ chains consisting of monomers with less resistance, ΔR, should exhibit more solitons statistically, at equal applied current, then chains with more such resistance.

It is demonstrated in the next section that relating the current-voltage quadratic relation problem to the d.c. current singularities of the long Josephson junction is straightforward, resulting in an explanation for the current-voltage relations of a collection of nonuniformly doped $(CH)_x$ chains, as both phenomena are based on the same soliton dynamics.

LONG JOSEPHSON JUNCTION ANALOGY

The problem of predicting the average voltage for a given applied current in a nonuniformly doped $(CH)_x$ chain is directly related to the explanation of the d.c. current singularities of the long Josephson junction (44-49) (Fig. 3C). The fluxon on the long Josephson junction is mathematically equivalent to a soliton and shares common analytic mechanics with the $(CH)_x$ soliton.

The sine-Gordon equation, without dissipative effects for which the soliton is a solution, is:

$$\frac{\partial^2\phi}{\partial x^2} - \frac{\partial^2\phi}{\partial t^2} = \sin\phi, \tag{3}$$

where ϕ is the phase difference between macroscopic quantum wave functions that characterize two super conductors and is directly analogous to the ϕ of the anharmonic oscillator model. The soliton (fluxon) solution to Equ. (3) is (50):

$$\phi = 4\tan^{-1}(f(x)g(t)) \tag{4}$$

and f and g are Jacobian elliptic functions defined (44-47):

$$\left(\frac{df}{dx}\right) = Af^4 + (1 + B)f^2 - C, \tag{5A}$$

$$\left(\frac{dg}{dx}\right) = Cg^4 + Bg^2 - A, \tag{5B}$$

where A, B and C are arbitrary constants. Taking $a = (1+B)/A$, $b = C/A$ for (5A) and $a = B/C$, $b = A/C$ for (5B), and differentiating, the form of Equ. (1) is obtained.

For the case of most interest, that with dissipative effects, we have:

$$\frac{\partial^2\phi}{\partial x^2} - \frac{\partial^2\phi}{\partial t^2} - \alpha[1+\varepsilon\cos\phi]\frac{\partial\phi}{\partial t} = \sin\phi, \tag{6}$$

where α and ε are constants. This form can be solved by assuming an x and t dependence of ϕ, e.g.:

$$\phi = \phi\,(x - ut) = \phi(\xi), \tag{7}$$

where u is a constant propagation velocity.

With α and ε related to an external energy source (46), Equ. (6) is:

$$\frac{\partial^2\phi}{\partial x^2} - \frac{\partial^2\phi}{\partial t^2} - \Gamma(1+\varepsilon\cos\phi)\left|\frac{\partial\phi}{\partial t}\right|\frac{\partial\phi}{\partial t} = \sin\phi-\gamma, \tag{8}$$

where $\Gamma = g_o\Phi_o/2\pi C$,

$\varepsilon = g_1/g_o$,

$\gamma = I_B/I_o$,

$\alpha = g_o|\nu|\,(\Phi_o/2\pi I_o C)^{1/2}$

and I_B is a uniformly distributed current bias for the long Josephson junction and a current bias for the $(CH)_x$ chain.

With this external source dependence, Equ. (8) has the desired quadratic current-voltage characteristic with normalized voltage (45):

$$v = \frac{1}{\pi k}\left(\frac{\gamma_o}{2}\right)^{1/2} \mathrm{dn}\left[\frac{1}{k}\left(\frac{\gamma_o}{2\Gamma u}\right)^{1/2}(\xi-\xi_o)\,;\,k\right] \tag{9}$$

where ξ_o is an arbitrary constant:

$$\xi = (x - ut);$$

and dn [.,.] represents a Jacobian elliptic function (51) of modulus k, $0<k<1$:

$$k = \left(\frac{2\gamma_o}{\gamma+\gamma_o}\right)^{1/2}; \tag{10}$$

$$\gamma = 2\Gamma u^2\,[(1-u^2)^2 + 4\Gamma^2u^4]^{1/2}. \tag{11}$$

Because

$$\frac{1}{k}\left(\frac{\gamma_0}{2\,u^2}\right)^{1/2} = 2nK(k), \tag{12}$$

(where ℓ is the length of the long Josephson junction (or $(CH)_x$ chain); K(k) is the complete elliptic integral of the first kind; and n is the number of solitons (fluxons) contained on it), the average value of voltage, v, is:

$$\langle v\rangle = nu/\ell \ . \tag{13}$$

As the number of solitons, n, will vary from monomer to monomer at constant applied current, depending on the doping condition, the average $\langle n\rangle$ is substituted for n and:

$$\langle v\rangle = \langle n\rangle u/\ell \ . \tag{14}$$

The current-voltage relations calculated with Equ. (14) are shown in Figure 3D. The similarity between the relations shown in Figure 3D, the putative relations of Figure 3B and the experimentally obtained relations of Figure 1, indicate that a possible explanation of the nonlinearity of Figure 1 is obtained.

CONCLUSION

1. Proteins and DNA are presently being considered as carriers of solitons of the *vibrational* type.

2. Doped trans-polyacetylene is a carrier of the *charge-domain-wall* type soliton and has been more intensively studied experimentally and theoretically with respect to soliton mediation.

3. At this time, *models* of soliton propagation offer a more rigorous *conceptual* platform.

4. The *anharmonic oscillator model* and the *fluxon* (or soliton) of the *long Josephson junction* offer a *conceptual* platform for examination of molecular solitons.

5. There is a remarkable similarity between the current-voltage relations of *inhomogeneously doped* trans polyacetylene and the current voltage characteristics of the long Josephson junction *in the presence of an external magnetic field applied transversely*, which can be explained theoretically.

6. The anharmonic oscillator model, fundamental to moving wave phenomena including the fluxon of the long Josephson junction, is thus advocated as *conceptual* platform for the study of molecular solitons.

REFERENCES

1. Zabusky, N.J. and Druskal, M.D., Phys. Rev. Lett. 18 (1965).
2. Zabusky, N.J., J. Comput. Phys. 43, 195 (1981).
3. Russell, Hohn Scott, "Report on Waves", British Association Reports (1834).
4. Davydov, A.S., J. Theoret. Biol. 66, 379 (1977).
5. Davydov, A.S., Phys. Stat. Sol. (b)90, 457 (1978).
6. Davydov, A.S., Int. J. Quant. Chem. 16, 5 (1979).
7. Davydov, A.S., Sov. Phys. JETP 51, 397 (1980).
8. Davydov, A.S., Phys. Stat. Sol. (b)102, 275.
9. Scott, A.C., Phys. Lett. 86A, 60 (1981).
10. Hyman, J.M., McLaughlin, D.W. and Seoh, A.C., Physica 3D, 23 (1981).
11. Balanovski, E. and Beaconsfield, P., Phys. Rev. Lett. 93A, 52 (1982).
12. Su, W.P., Schrieffer, J.R. and Heeger, A.J., Phys. Rev. Lett. 42, 1698 (1979).
13. Su, W.P. Schrieffer, J.R. and Heeger, A.J., Phys. Rev. B22, 2099 (1980).
14. Heeger, A.J. and MacDiarmid, A.G., in: "Physics in One Dimension", Bernasconi, J. and Schneider, T. (eds.), Springer, NY (1981).
15. Scott, A.C., in: "Nonlinear Phenomena in Physics and Biology", Enns, R.H., Jones, B.L., Miura, R.M. and Rangnekkar, S.S. (eds.), Plenum, pp. 179-193, (1981).
16. Shirakawa, H., Ito, T. and Ikeda, S., Makromol., Chemie 179, 1565 (1978).
17. Park, Y.W., Heeger, A.J., Druy, M.A. and MacDiarmid, A.G., J. Chem. Phys. 73, 946 (1980).
18. Epstein, A.J., Gibson, H.W., Chaikin, P.M., Clark, W.G., and Gruner, G., Phys. Rev. Lett. 45, 1730 (1980).
19. Roland, M., Bernier, P., Lefrant, S., and Aldissi, M., Polymer 21, 1111 (1980).
20. Gould, C.M., Bates, D.M., Bozler, H.M., Heeger, A.J., Druy, M.A. and MacDiarmid, A.G., Phys. Rev. B23, 6820 (1981).
21. Moses, D., Chen, J., Denenstein, A., Kaveh, M., Chung, T.C., Heeger, A.J., MacDiarmid, A.G., and Park, Y.W., Solid State Comm. 40, 1007 (1981).
22. Chiang, C.K. and Franklin, A.D., Solid State Comm., 40, 775 (1981).
23. Yamamoto, Y., Yoshino, K. and Inuishi, Y., Japan J. Appl. Phys. 19, 991 (1980).
24. Ikehata, S., Kaufer, J., Woerner, T., Pron, A., Druy, M.A., Sivak, A., Heeger, A.J. and MacDiarmid, A.G., Phys. Rev. Lett. 45, 1123 (1980).
25. Chien, J.C.W., Wneck, G.E., Karasz, F.E., Warakomski, J.M., Dickenson, L.C., Heeger, A.J. and MacDiarmid, A.G., Macromolecules 15, 614 (1982).

26. Kivelson, S., Mol. Cryst. Liq. Cryst. 77, 65 (1981).
27. Kivelson, S., Phys. Rev. Lett. 46, 1344 (1981).
28. Nechtschein, M., Devreux, F., Greene, R.L., Clarke, T.C. and Street, G.B. Phys. Rev. Lett. 44, 356 (1980).
29. Chen, C.K., Park, Y.W., Heeger, A.J., Shirakawa, H., Louis, E.J. and MacDiarmid, A.G., J. Chem. Phys. 69, 5098 (1978).
30. Barrett, T.W., Naval Air Development Center Memoranudm Report, in press (1983).
31. Bruce, A.D., in: "Solitons and Condensed Matter Physics", A.R. Bishop and T. Schneider (eds.), Springer, NY (1978).
32. Burce, A.D. and Cowley, R.A., Adv. in Physics 29, 219 (1980).
33. Murata, K.K., Phys. Rev. B11, 462 (1975).
34. Aubry, S., J. Chem. Phys. 62, 3217 (1975).
35. Aubry, S., J. Chem. Phys. 64, 3392 (1976).
36. McMillan, W.L. Phys. Rev. B14, 1496 (1976).
37. Hohenberg, P.C. and Halperin, B.I., Rev. Mod. Phys. 49, 523 (1977).
38. Halperin, B.I., Hohenberg, P.C. and Ma, S., Phys. Rev. B10, 139 (1974).
39. Halperin, B.I., Hohenberg, P.C. and Ma, S., Phys. Rev. B13, 4119 (1976).
40. Schneider, T. and Stoll, E., Phys. Rev. Lett. 31, 1254 (1973).
41. Schneider, T. and Stoll, E., Phys. Rev. Lett. 35, 296 (1975).
42. Schneider, T. and Stoll, E., Phys. Rev. 13, 1216 (1976).
43. Schneider, T. and Stoll, E., Phys. Rev. 17, 1302 (1978).
44. Chen, J.T., Finnegan, T.F. and Langenberg, D.N., Physica 55, 413 (1971).
45. Parmentier, R.D. in: "Solitons in Action", K. Lonngren and A. Scott (eds.) Academic, New York, pp. 173-199 (1978).
46. Parmentier, R.D. and Costabile, G., Rocky Mountain J. Math. 8, 117 (1978).
47. Costabile, G., Parmentier, R.D., McLaughlin, D.W. and Scott, A.C., Appl. Phys. Lett., 32, 587 (1978).
48. Costabile, G., Cucolo, A.M., Pace, S., Parmentier, R.D., Savo, B. and Vaglio, R., in: "SQUID '80 - Superconducting Inter ference Devices and Their Applications", H.D. Hahlbohm and H. Lubbig (eds.), Walter de Gruyter, Berlin, pp. 147-155 (1980).
49. Costabile, G. and Parmentier, R.D., Lett. Nuovo Cimento 27, 270 (1980).
50. Lamb, G.L., Rev. Mod. Phys. 43, 99 (1971).
51. Byrd, P.F. and Friedman, M.D., "Handbook of Elliptic Integrals for Engineers and Physicists", Springer, New York (1954).
52. Fulton, T.A. and Dynes, R.C., Solid State Comm. 12, 57 (1973).

NONLINEAR EXCITATIONS IN POLYACETYLENE

W. P. Su

Department of Physics
University of Illinois at Urbana-Champaign
Urbana, IL 61801 USA

As one of the simplest conjugated systems the polyacetylene chain has been of interest to chemists for a long time.[1] In recent years it has attracted physicists' attention as a quasi one dimensional material. It is a highly nonlinear system. Being very interesting from many points of view, it serves to illustrate the fruitfulness of interdisciplinary interactions.

Most of the quantitative results have been obtained using a very simple tight-binding type Hamiltonian.[2,3] The Hamiltonian is linear in the electron-phonon coupling. Within the adiabatic approximation it is effectively a Huckel type theory.

The balance of the electronic energy and the lattice strain energy leads to many nonlinear solutions. For example it is shown that the ground state configuration of a polyene chain has a bond length modulation - single and double bonds alternate along the chain. Although the result is well known, the definition of a single or double bond is different from that which is adopted in elementary chemistry. The double bond here means a shorter length than a single bond and doesn't imply a pair of localized π-electrons.

For an infinite chain there are two degenerate ground states: the A phase and the B phase. B is related to A by interchanging the double and single bonds. In either phase there is an electronic gap of about 1.5 eV separating the filled valence band from the empty conduction band. A kink that connects a segment of the chain in the A phase to another segment in the B phase is a topological soliton. The disruption of the perfect single-double bond alternation pattern near a kink also introduces a distortion in the electronic spectrum. There is always a localized molecular orbital within the gap. The additional optical absorption caused by the gap state provides a most direct way of observing the kinks.[4,5] The spin density associated with the gap state can also be scrutinized by all sorts of sophisticated electron spin resonance and nuclear magnetic resonance experiments.[6-8] Most of the experimental results confirm the theoretical predictions that the kink has a width of about 10 lattice constants and has a very small effective mass (a few electronic masses).[9]

Neutral kinks occur naturally in finite odd chains.[10] They carry a net spin but no charge. The separation of charge and spin in a particle-like entity (the kink) is something that physicists find fascinating.[11]

Charged but spinless kinks can be produced by doping[12] or by photoexcitation.[13-15] The photoinduced absorption spectrum has the same features as the dopant induced absorption spectrum, indicating that these features are intrinsic to the kinks.[16]

The dynamical behavior of a single soliton and a soliton-antisoliton pair is only beginning to be explored.[17] Besides the kinks, there are also the polaron solutions[18] as discussed by Dr. Bishop in this meeting and possible others as well.

For an elementary survey the reader is referred to the forthcoming review articles.[19,20]

This work was supported by the Material Research Laboratory under National Science Foundation grant NSF DMR 81-17182.

REFERENCES

1. L. Salem: The Molecular Orbital Theory of Conjugated Systems, W. A. Benjamin, New York 1966.
2. W. P. Su, J. R. Schrieffer and A. J. Heeger, Phys. Rev. Lett. 42, 1698 (1979).
3. W. P. Su, J. R. Schrieffer and A. J. Heeger, Phys. Rev. B 22, 2099 (1980).
4. N. Suzuki, M. Ozaki, S. Etemad, A. J. Heeger and A. G. MacDiarmid, Phys. Rev. Lett. 45, 1463 (1980).
5. A. Feldblum, J. H. Kaufman, S. Etemad, A. J. Heeger, T. C. Chung, and A. G. MacDiarmid, Phys. Rev. B, 26, 815 (1982).
6. B. R. Weinberger, E. Ehrenfreund, A. Pron, A. J. Heeger, and A. G. MacDiarmid, J. Chem. Phys. 72, 4749 (1980).
7. M. Nechtschein, F. Devreux, F. Genoud, M. Guzlielmi, and K. Holczer, phys. Rev. B 27, 61 (1983).
8. H. Thomann, L. R. Dalton, Y. Tomkiwicz, N. S. Shiren, and T. C. Clarke, Phys. Rev. Lett. 50, 533 (1983).

9. Z. Vardeny, J. Orenstein, and G. L. Baker, Phys. Rev. Lett. 50, 2032 (1983).
10. W. P. Su, Solid State Commun. 35, 899 (1980).
11. W. P. Su and J. R. Schrieffer, Phys. Rev. Lett. 46, 738 (1981).
12. H. S. Ikehata, J. Kaufer, T. Woerner, A. Pron, M. A. Druy, A. Sivak, A. J. Heeger, and A. G. MacDiarmid, Phys. Rev. Lett. 45, 1123 (1980).
13. W. P. Su and J. R. Schrieffer, Proc. Natl. Acad. Sci. USA, 77, 5626 (1980).
14. J. Orenstein and G. L. Baker, Phys. Rev. Lett. 49, 1043 (1980).
15. J. D. Flood, E. Ehrenfreund, A. J. Heeger, and A. G. MacDiarmid, Solid State Commun. 44, 1055 (1982).
16. G. B. Blanchet, C. R. Fincher, T. C. Chung, and A. J. Heeger, Phys. Rev. Lett. 50, 1938 (1983).
17. W. P. Su, Mol. Cryst. Liq. Cryst. 77, 265 (1981).
18. D. K. Campbell and A. R. Bishop, Phys. Rev. B24, 4859 (1981).
19. W. P. Su, "Solitons in Polyacetylene - A Review", to appear in *Handbook on Conducting Polymers*, edited by T. Skotheim, Marcel Dekker, New York.
20. J. R. Schrieffer and A. J. Heeger, Reviews of Modern Physics.

PROSPECTS FOR A BIOELECTRONIC TECHNOLOGY

COMMUNICATION WITH SUBMICRON STRUCTURES: PERSPECTIVES IN THE APPLICATION OF BIOMOLECULES TO COMPUTER TECHNOLOGY

Albert F. Lawrence and Robert R. Birge

Hughes Aircraft Company, PO Box 9399, Long Beach, CA, and
University of California, Chemistry Dept., Riverside, CA

1. EVOLUTION OF ELECTRONIC DEVICES

That demands on computer technology keep several steps ahead of the hardware is proverbial. The needs of the scientific and engineering communities for numerical modeling and image processing have stimulated an exponential growth in processor speeds and memory capacities. This growth has been sustained by advances in device physics since the early 1950's. The major portion of these advances may be attributed to a steady decrease in circuit dimensions. The limits to miniaturization in a semiconductor-based technology are now in sight. The breakdown of transport laws in bulk semiconductor samples with dimensions below 0.1 micron precludes use of conventional materials and structures for devices at and below the 0.1 micron scale. In order to sustain further progress, circuit technology must turn to alternative materials.

The fundamental limits of computer speed are determined by gate and communication delays. From the standpoint of fundamental physics, the speed of communication between circuits is limited by special relativity, while the speed and scale of individual circuits are bound by quantum mechanics. As participants in this conference have shown, analogs of gates, valves, memory elements, oscillators and clocks all exist at the molecular level (See Carter, these proceedings). Thus, on the scale of individual circuits, the evolution of computers toward higher speeds, greater circuit density and lower power consumption presently seems limited only by fundamental molecular physics. As an illustration of this point, Figure 1 compares resolution limits for various lithographic techniques against molecular dimensions of organic polymers and dyes. Electron-beam (EB) techniques may be used to build structures compatible in

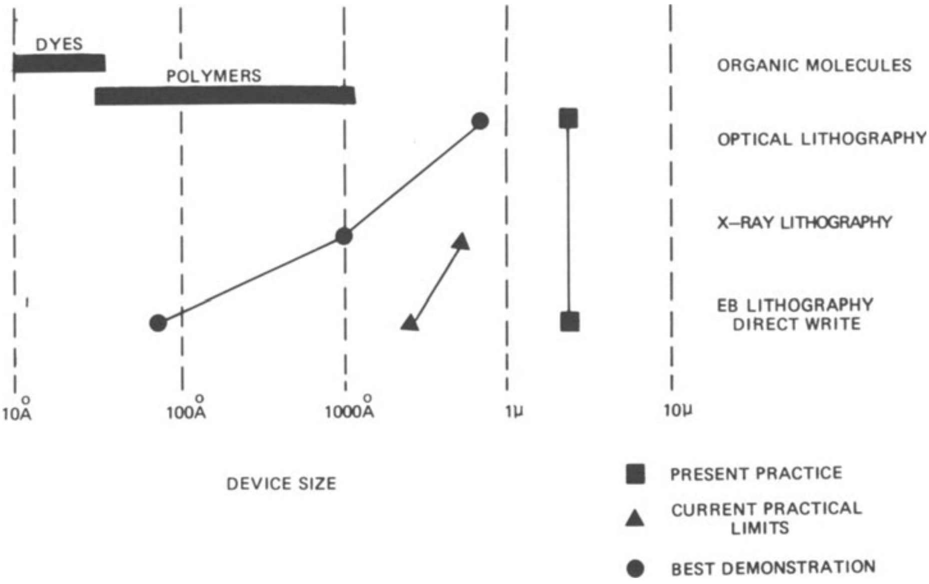


Figure 1. Resolution limits for various thin-film fabrication techniques compared to molecular dimensions.

size with polymers or biological macromolecules.[1] Since suitable techniques could be developed for the synthesis and attachment of molecular structures to a proper substrate, we have at hand many of the techniques necessary for the fabrication of circuits on the nanometer scale. McAlear, (these proceedings) has discussed the issue of synthesis from a biological perspective.

The principal advantage of nanometer scale structure is circuit density. Assuming resolution limits consistent with current (EB) technology,[1] we would be able to pack as many as a hundred thousand times more circuits onto a planar chip than is the current practice. Development of molecular-level technology would also pay enormous dividends in circuit speeds. This is because higher densities imply reduced signal propagation time between individual circuits. In addition, typical time scales for molecular processes are in the picosecond range or faster. Molecular-level circuits would offer increases in speed three or more orders of magnitude beyond current silicon technology.

Organic polymers and dyes possess a wide range of desirable optical and electrical properties. The first step toward the molecular scale is the use of very thin films. Development and exploitation of techniques for deposition of monomolecular films is an active area of research. Thompson (these proceedings) has described the use of very thin films as tunnel barriers in electronic devices.

Langmuir-Blodgett deposition, chemisorption, and surface polymerization techniques have all been used to form very thin films. In addition to being compatible with present thin-film techniques, monomolecular films offer the possibility of preserving the useful macroscopic properties in two dimensions, while at the same time utilizing the microscopic properties in the third. Two examples, which we will discuss in this presentation, are thin film optical memory devices and active tunnel barriers.

2. PRESENT LIMITS TO COMMUNICATION WITH SUBMICRON STRUCTURES

Assuming that synthesis and fabrication are possible, communication is a major practical problem with nanometer structures. Building an assembly of nanometer circuit elements loses some of its intrinsic appeal if input and output require a roomful of high-precision electronics. Solving the I/O problem forces us to consider means of communication through intermediate structures. Use of monomolecular films offers the possibility of fabrication of the necessary intermediate structures.

Two other technical issues are compatibility of power dissipation levels (energy required to switch a gate) and switching times. Figure 2 places these issues in the perspective of current technology. This figure compares power dissipation levels and switching times for several different semiconductors, Josephson junctions and molecules which exhibit switching behavior. As one might infer from the figure, Josephson junction technology bridges the gap between characteristic times for semiconductor gates and characteristic times for molecular transitions.[2,3,4] In addition, Josephson junction energy levels are intermediate between energies necessary to switch molecular states and energies necessary to switch semiconductor gates. The use of Josephson junction devices as intermediate between molecular gates and the macroscopic world is a reasonable prospect. An increasing amount of semiconductor circuitry for sensors is being run at cryogenic temperatures. General considerations of power dissipation per unit volume versus circuit density seem to require the low power dissipation levels provided by cryogenic circuitry. Josephson junction technology meets the criteria of speed and low power dissipation. Finally, because Josephson junctions require dielectric barriers whose thicknesses are measured in nanometers,[2] this technology offers a convenient synergy with thin film technology.

3. INVESTIGATION OF MOLECULAR PROCESSES IN THIN FILMS

Before turning to discussion of some potential applications of biomolecules to computer technology, we will discuss a few techniques for investigation of molecular processes in thin films.

Present laser techniques permit focussing of laser beams (in the visible range) to a one micron spot. Various techniques are also

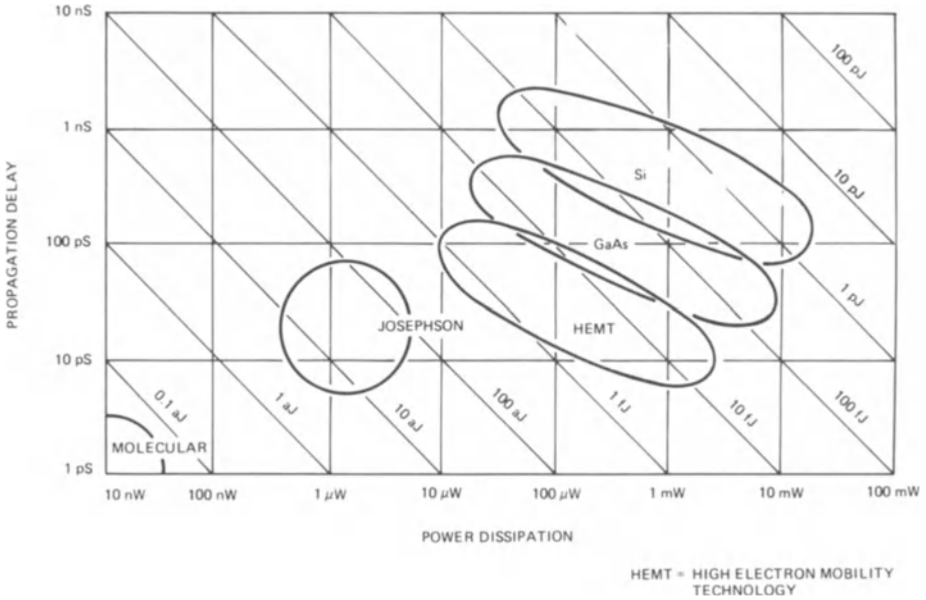


Figure 2. Gate delays and power dissipation levels for solid state devices compared to transition times and power consumption levels for molecular processes.

available to obtain light pulses of one picosecond or less in duration. The combination of small spots and short pulses is obtainable with off-the-shelf equipment. Although resolution limits imposed by the wave nature of light do not permit optical addressing of submicron structures, laser techniques are increasingly useful in investigating the dynamics of molecules in small structures.

We will now turn our attention to techniques which may be used to investigate the low-energy end of the electromagnetic spectrum. Frequencies corresponding to acoustic phonons are of particular interest. Acoustic phonons corresponding to the 1 to 200 wavenumber range (30 gigahertz to 6 terahertz) play a major role in device physics. Three aspects of the acoustic phonons seem to be particularly relevant to performance of small devices. First, acoustic phonons determine the upper limits of device speeds. This is certainly the case in superconductive electronics, where the formation of Cooper pairs, and thus the onset of the superconductive state is due to electron-phonon interactions.[2] Second, several molecular switching processes occur in the picosecond range.[5,6] Applications of molecular processes in electronic devices must take account of possible resonances between molecular oscillations and oscillations occurring at the characteristic frequencies of the device. Third, acoustic phonons are intimately connected with solitary wave processes in long chain polymers and molecular crystals.[7,8] The so-called solitonic processes seem to offer a means of energy transfer (communication) between molecular structures which is not subject to radiative processes and is exceptionally stable against dissipation by non-radiative processes.[7] Detection of characteristic phonon signatures may be a great aid in the study of solitary wave phenomena in molecules.

A number of critical issues in the development of electronics applications of biomolecules turn on the possible interactions of far infrared and millimeter wave frequencies in submicron structures. It is interesting to note the parallel between this observation and the biologist's increasing appreciation of the significance of millimeter and submillimeter frequencies in the molecular processes of living cells. Techniques which can be applied to the study of molecular resonances in artificial thin films may also provide information useful to the biologist.

The simplest method for studying low-frequency molecular vibrations in thin films involves the measurement of tunneling currents at low temperatures in metal-insulator-metal (M-I-M) junctions or superconductor-insulator-superconductor (S-I-S) junctions.[9] Analysis of current-voltage relations in M-I-M junctions is the basis of the technique of inelastic electron tunneling spectroscopy (IETS). Interaction of tunneling electrons with excitations in the barrier (molecular vibrations, excitons, solitons, etc.) provides extra channels to a tunneling electron. This is because an electron losing energy to a barrier excitation may tunnel to a lower energy level in the counter electrode. Raising the relative voltage between source and counter electrodes makes more of these lower energy channels available to the tunneling electron.[9] This is reflected in the I-V curves for the junction, as shown in Figure 3.

At the relative voltage a tunneling channel becomes available, the I-V curve shows a bend, the first derivative (with respect to voltage) shows a step and the second derivative shows a spike. Spikes in the second derivative of the I-V curve reflect lines in the inelastic tunneling spectrum of the barrier. These lines are a combination of the usual infrared absorption and Raman spectra of the molecules in the barrier; with a geometric selection rule.[9] Generally, tunneling electrons interact with molecular dipoles oriented perpedicularly to the barrier.

IETS generally requires the measurement of the second harmonic of the voltage response to a sinusoidally modulated current source.[9] This makes the technique rather slow. Sensitivity is quite high, however. Use of a differential bridge permits measurement of the spectrum of a sample of as few as 10 billion molecules in a film several square millimeters in area.[9] Scaling present experimental

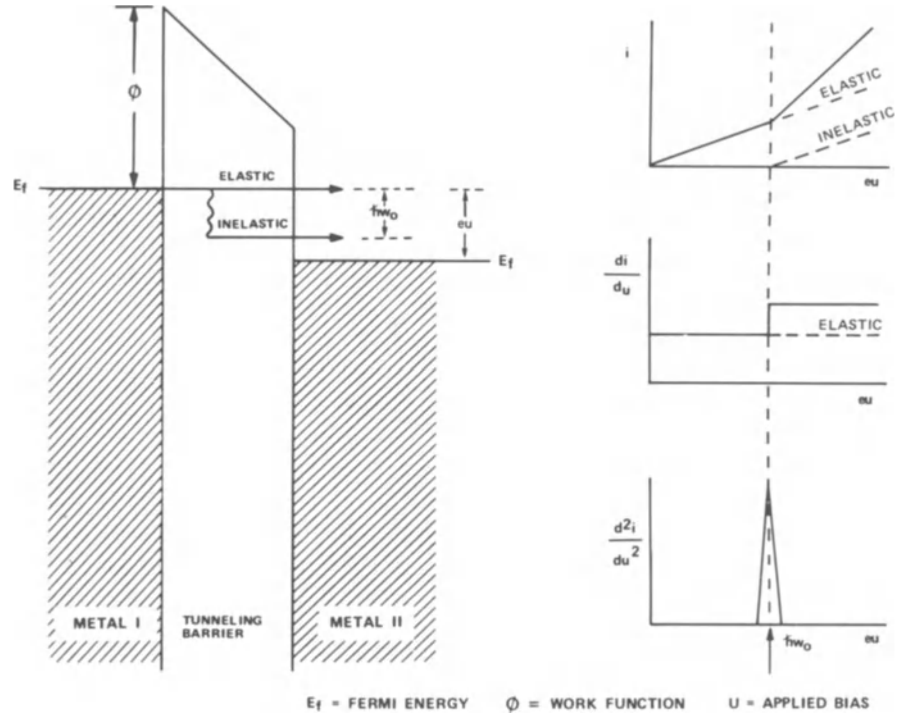


Figure 3. (a) Energy level diagram for elastic and inelastic tunneling processes. Inelastic tunneling channels are opened when a tunneling electron is permitted to give its energy to molecular vibrations in the barrier layer. (b) The characteristic energy losses due to inelastic scattering interactions may be determined from the second derivative of the I-V curves for the tunnel device.

apparatus down to micron dimensions and using high sensitivity, low noise cryogenic electronics should improve both speed and sensitivity several orders of magnitude.

The Josephson effect provides more promising techniques from the standpoint of speed. At least two different approaches might be taken toward the study of the dynamics of polymers and dyes in the solid state. The first would be to investigate interactions between quasi-particle tunneling and the Josephson current in a Josephson junction. Organic molecules in the barrier region would interact with the quasi-particle tunneling. Changes in the quasi-particle tunneling current couple to the Josephson current near the Reidel singularity.[2] Relevant junction characteristics might be measured in a SQUID configuration.

The second technique would use a SQUID device to measure magnetic transitions in a population of molecules. A simple calculation shows that a SQUID magnetometer operating at the present limits of sensitivity could detect field changes which are due to the change induced by 1000 simultaneous spin flips in a small material sample. Measurements could be made either indirectly, as changes in magnetic susceptibility, or directly, as an induced magnetic field. Reported measurements of magnetic susceptibility approach the theoretical limits of sensitivity.[2] Unfortunately, indirect measurements as presently done by SQUID devices are rather slow, but direct measurements could be made on a picosecond time scale. The penalty for making the fastest possible measurement is that magnetic field changes would have to be on the order of a single flux quantum. This means that the magnitude of the effect to be detected would have to be several orders of magnitude greater than a thousand spin flips. This puts direct magnetic measurements with SQUIDS in the same sensitivity range as IETS, but the speed is at least 10 orders of magnitude greater than IETS as presently practiced in the laboratory.

4. PROSPECTS FOR DEVICES BASED ON ORGANIC OR BIOMOLECULES

Present prospects for devices at the molecular scale can be guaged by the present limits of laboratory measurement by electronic techniques, which are three or four orders of magnitude away from detection of events in a single molecule. This compares rather unfavorably with the sensitivities of biological systems reported at this conference. It would seem that the device physicist could learn from nature.

Optical techniques may be somewhat better than electronic for detecting events in small ensembles of molecules, but the energies necessary to produce photons for the detectors seem to be incompatible with power dissipation limits in dense structures. Although appropriate for single experiments and simple devices, optical circuits are not useful for the very large scale integration which is coming into use for the present generation of computers.

One solution to the problems of measurement is to integrate experiment and detectors into microcircuits on the same chip. This would aid progress toward molecular scale devices. In addition to producing devices with useful properties, thin film techniques may also lead to some interesting experiments. One class of proposed devices is Josephson junctions with optically active barrier layers. We discuss this idea as a particularly interesting example of the class of devices based on monomolecular films below.

One may also envision a solution which evades the problems of scale and power dissipation in optical devices. The basic paradigm is to use molecular transitions as a type of content addressable memory.[10] Content addressable memories are advantageous for a number of computer applications.[11]

For a content addressable memory based on optical spectra, we require a material composed of molecules with several stable states, each with a distinct set of absorption lines. Absorption of a photon would induce a change of state and a corresponding change in the absorption spectrum. Given that the molecule possesses spectral lines which may independently register the absorption of photons at n different wavelengths and that there are no interference effects among the different spectral changes, that molecule would serve as an n-bit memory.

Memory storage in a material composed of these idealized molecules would involve changing a small sub-population of the molecules in the material. Ideally, it would not be necessary to write to any particular spatial location, so organization of the material on a molecular level would not be necessary. Rather, memory capacity would correspond to the ratio of sample size to the minimal population for which a change may be detected.

Although no materials with three or more homogeneous independent states are known to the authors, even a two-state memory molecule would be extremely useful. We discuss a device based on such a substance in the last section.

The most serious obstacle to designing a molecular level digital computer is that a sufficiently accurate theory of the dynamics of large molecules is not yet available. We cannot hope to specify a complex structure like a digital computer until we have a means of embodying its fundamental operations in an actual physical structure. Although models of quantum-level Turing machines have been proposed[12], none of these models exhibit realistic molecular forces in their Hamiltonian formulation.[13] The appropriate extension of present models should provide several key features:

(1) a library of molecular structures which embody the fundamental operations of combinatorial logic,

(2) molecular control structures which may be coupled to the computational elements to force unambiguous sequences of operations,

(3) dynamical models of control and logical elements and their interactions, and

(4) models of scattering events which would provide input and output operations.

Such a theory would point the way toward synthetic molecular structures capable of executing stored programs. Because of the requirement that the results of the operation of this computer be unambiguous, the observables would more likely be classical rather than quantum in nature. This may be accomplished by statistic means, i.e., duplication of structure and operations, or through dynamical processes which carry a large momentum. Because solitonic processes carry a relatively large momentum and may be modeled semiclassically, this has led to the speculation that solitons are the appropriate dynamic vehicles for computer operation (Carter, these proceedings). The complexities of realistic models of the types of nonlinear molecular dynamics which might produce solitonic events (Scott and Lomdahl, these proceedings) may make numerical simulation by present computers impossible. At this point, the requirement for more powerful computers becomes circular. We need better computers in order to build better computers.

5. OPTICALLY SENSITIVE JOSEPHSON JUNCTIONS

Given presently-available technology, the optimal computer should probably be a hybrid. Josephson junction technology seems to offer the fastest possible logic with the lowest power dissipation. Optical devices, on the other hand, seem to offer the highest memory capacities. Building a hybrid of optical and Josephson junction components is impeded by the lack of a practical optical-Josephson junction interface.

Interactions with organic materials are well known for non-Josephson currents of tunneling electrons. Laboratory demonstrations include a reported Josephson interaction with protoporphyrin IX Fe (+++) in a point-contact junction.[14] Also, substances composed of similar heterocyclic molecules, such as the pthalocyanines, have been shown to behave as semiconductors. The conductivity of these substances may be rapidly modulated by exposure to circularly polarized light.[15] Given that either porphyrins or pthalocyanines maintain their optical activity down to temperatures necessary for operation of Josephson junctions, an optically sensitive junction may be built as illustrated in Figure 4.

Polyacetylene is another candidate for an optically active barrier material. Polyacetylene exhibits a photovoltaic effect. Just as in photovoltaic semiconductors, this is the result of the formation of an electron-hole pair subsequent to absorption of a photon by the polyacetylene molecule. Su and Schrieffer[16] have shown that the conjugated bond system of the molecule is unstable in the presence of an electron-hole pair and distorts to form a soliton-antisoliton pair. The time for formation of the two solitons is about 0.1 picosecond.[16] Furthermore, the energy barriers due to lattice discretization for a soliton moving along a polyacetylene molecule is about 0.004 electron volts.[17] This is in resonance with characteristic frequencies in Josephson junction devices.[2] A junction

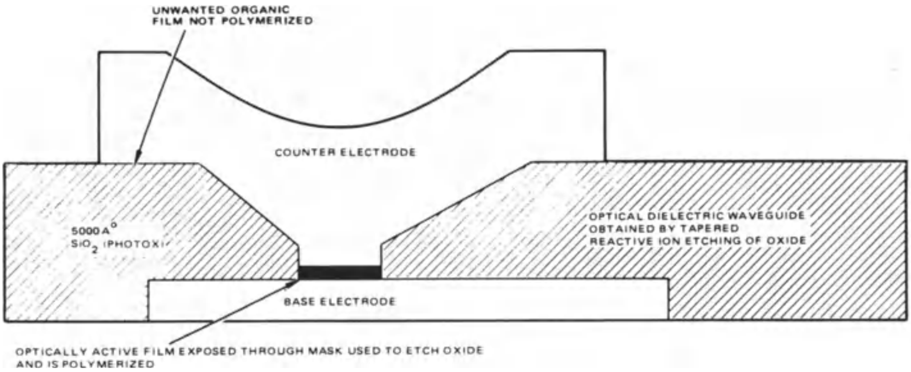

UNWANTED ORGANIC FILM NOT POLYMERIZED
COUNTER ELECTRODE
5000 A°
SiO_2 (PHOTOX)
OPTICAL DIELECTRIC WAVEGUIDE OBTAINED BY TAPERED REACTIVE ION ETCHING OF OXIDE
BASE ELECTRODE
OPTICALLY ACTIVE FILM EXPOSED THROUGH MASK USED TO ETCH OXIDE AND IS POLYMERIZED

Figure 4. Edge-coupled organic barrier in Josephson junction. Light interacting with molecules in the barrier changes electrical properties of the junction.

with a polyacetylene barrier region should show a strong response to photons of the proper energy. Such a device would act as a picosecond optical sensor.

6. A HIGH SPEED OPTICAL MEMORY BASED ON BACTERIORHODOPSIN

The second component of our proposed hybrid computer is a high-capacity optical memory. A thin layer of bacteriorhodopsin may be used as the active component of an optical memory device. The light-adapted form of bacteriorhodopsin (bR) has two, and only two, stable species of liquid nitrogen temperature (77K).[18] These two species are called BR and K where K is the informal name for the primary photoproduct. These two species can be photochemically interconverted with transformation times of approximately 10 picoseconds. Furthermore, these two species have distinct absorption maxima (568 nm for BR, 610 nm for K).[17] The high interconversion rates, the distinct absorption maxima, and the high photochemical stability of this protein make it an optimal biomolecular storage device. Furthermore, the quantum yields for the BR --> K (0.33) and the K --> BR (0.67) photoconversions are significantly higher than those of competitive molecular storage devices.[18] While this is not a necessary prerequisite for a device, it indicates that much lower optical fluxes are required to initiate bit flipping which ultimately lowers the heat generated during read/write cycling. Furthermore, the bit densities are many orders of magnitude higher than magnetic flux devices, which provides both intrinsic speed as well as storage capacity.

Highly stable dried films of bR may be deposited onto a quartz or sapphire substrate. Such a structure would comprise the active element of an optical disc storage device. Information would be stored by assigning BR to bit 0 and K to bit 1. Irradiation at 500 nm converts BR to K with a quantum efficiency of 0.33. Resetting the bit is accomplished using 633 nm laser radiation which converts K to BR with a quantum efficiency of 0.67. The read operation is performed using both lasers in a coupled mode with the 633 nm beam reading the track, and 1 bit being reset immediately with the 500 nm beam. This technique is the optical analog of microelectronic refresh techniques.

Studies of cell membranes containing bR indicate that three bR molecules occupy a unit cell with dimensions of approximately 50 x 50 Angstroms. If we arbitrarily assign a single bit to the unit cell (three chromophores/bit), a single bit occupies an area of only 2.5×10^{-13} cm^2. The active surface area on a small 5.25" disc is approximately 400 sq. cm. Accordingly, a single small disc has the potential of storing 200,000,000 megabytes. This is, of course, an unrealistic number for present technology because it is virtually impossible to focus a laser beam onto such a small area. However, if we include Gaussian optical limitations in our calculations, and use the 400 sq. cm area, our disc surface could easily store 500 megabytes and possibly an order of magnitude more. Devices of this type are currently being investigated at the University of California, Riverside.

ACKNOWLEDGMENT

This work was supported by Office of Naval Research Contract Number N0001483-C-0010.

REFERENCES

1. G. R. Brewer, ed., "Electron-beam Technology in Microelectronic Fabrication," Academic Press, New York (1980).

2. A. Barone and G. Paterno, "Physics & Applications of the Josephson Effect)," Wilery-Interscience, New York (1982).

3. G. Rabbat, ed., "Hardware and Software Concepts in VLSI," Van Nostrand-Reinhold, New York (1983).

4. A. Laubereau and W. Kaiser, Rev. Mod. Phys., 50:507 (1978).

5. R. R. Alfano, ed., "Biological Events Probed by Ultrafast Laser Spectroscopy," Academic Press, New York (1982)

6. R. R. Birge, Annu. Rev. Biophys., Bioeng, 10:315 (1981).

7. A. S. Davydov, Physica Scripta., 20:387 (1979).

8. W. P. Su, J. R.Schrieffer and A. J.Heeger, Phys. Rev. B, 22:2099 (1980).

9. P. K. Hansma, ed., "Tunneling Spectroscopy," Plenum Press, New York (1982)

10. A. F. Lawrence, "How Do We Talk to Nanometer Structures", in Proceedings of the 2nd International Conference in Molecular Electronics, Washington, D.C., (1983, To Appear).

11. C. Foster, "Content Addressable Parallel Processors," VanNostrand-Reinhold, New York (1977).

12. P. A. Benioff, Intl. Jour. Theor. Phys., 21:177 (1982).

13. H. F. Hameka, "Advanced Quantum Chemistry," Addison-Wesley, New York (1965).

14. M. Puma and B. S. Deaver, Jr., Appl. Phys. Lett., 19:539 (1971)

15. T. W. Barrett, Physics Letters, 91A:139, (1982), Physics Letters, 92A:309, (1982.

16. W. P. Su and J. R. Schrieffer, Proc. Natl. Acad. Sci., USA, 77:5626 (1980).

17. D. Baeriswyl, G. Harbeke, H. Keiss, & W. Meyer, Polyacetylene, in "Electronic Properties of Polymers," J. Mort and G. Pfister, eds., Wiley-Interscience, New York (1982).

18. R. R. Birge and B. M. Pierce, "The Nature of the Primary Photochemical Event in Bacteriorhodopsin and Rhodopsin," in Proceedings of the International Conference on Photochemistry and Photobiology, Alexandria, Egypt, (1983 To Appear).

LANGMUIR-BLODGETT FILMS IN MICROELECTRONICS

E. D. Thompson

Computer Science and Electrical Engineering
Lehigh University
Bethlehem, PA 18015, U.S.A.

ABSTRACT

The potential application of Langmuir-Blodgett-deposited films in microelectronics is discussed. The application areas which appear most promising are those requiring near perfect and uniformly thick films in the thickness range 2.5-50 nm. Three areas are discussed: the use of Langmuir-Blodgett films as electron resists in submicron device fabrication; the use of these films as insulators in metal-insulator-semiconductor (MIS) structures in which the semiconductor is a compound one; and the use of Langmuir-Blodgett films as tunneling barriers in Josephson junction electronic circuits. It is pointed out that, while preliminary evidence indicates high potential payoff for these films, the amount of data and research is very limited.

I. Introduction

Langmuir-Blodgett-deposited films[1] have, to date, been primarily the study focus of physical chemists and physicists who have been concerned about the thermodynamics and the phase transitions of the almost-ideal two dimensional substances. There now appears to be an increasing interest in the application of these films to advancing the state of microelectronics.[2] It is the purpose of this paper to review some recent encouraging results in the applications area. It is pointed out at the outset, however, that much remains to be learned about the electrodynamic, thermal, and mechanical properties of these very interesting films before devices can be materialized. Because of the close relationship between the structure and properties of some Langmuir-Blodgett films and biological substances, understanding the electrodynamics of Langmuir-Blodgett films may assist in providing a deeper understanding of the dynamical processes occurring within biological materials.

Langmuir-Blodgett films appear attractive in applications requiring almost perfect films and also very thin films, films of the order of 2.5-50 nm thickness. The thickness of such multilayer films is determined by the number of depositing operations performed, countable integers, and the thickness of each layer. The latter is determined by the chemistry of the starting materials.

This thickness dependence contrasts sharply with that of the more normal and usual processing procedures in microelectronics such as spin-and-spray-deposition, sputtering, evaporation, and chemical vapor deposition. In all these procedures, the thickness depends upon the physical process variables and it is a credit to this industry that such close, but expensive, control can be maintained on these variables.

Secondly, the Langmuir-Blodgett films appear attractive as insulating and passivating layers on compound semiconductors. The native oxides of these materials have not proven to be effective layers and alternative materials are being explored, most commonly vapor- and sputter-deposited SiO_x. Electrical stability and low interface densities are the important requirements in the application rather than the thickness uniformity and control.

The details of the Langmuir-Blodgett process are well described in reference 1, but a brief summary is given here. The process is the transfer of a floating film to a substrate which is mechanically transported across the liquid-gas interface. The transport may be in either direction, and one layer of the film is deposited with each transport. The film, such as barium stearate, vinyl stearate, or diacetylene, consists of chain-like molecules the length of which depends upon the number of links or cells in the aliphatic chain and determines the film thickness. The film is placed upon the liquid sub-phase, usually water, by first dissolving the film's starting material in a solvent such as heptane, depositing the solution onto the sub-phase surface with a syringe, and then allowing solvent evaporation. Before transference to the substrate, the film is compressed into a close-packed, two-dimensional, molecular array with the hydrophobic molecular tails away from the sub-phase and the hydrophilic molecular heads next to the sub-phase, and this compressional force is maintained constant throughout the transference by means of a feedback control loop. While the process is thus simply described, in fact, the successful carrying out of the deposition requires great attention to mechanical details, to high-purity starting materials and sub-phase, and to control over all deposition steps.

Having indicated the essentials of the Langmuir-Blodgett process and the primary motivations for the use of such films in microelectronics, we shall now proceed to describe recent results in three application areas. The first of these is the use of such films as electron resists for sub-micron circuit fabrication. As feature sizes shrink from the current micrometer dimemsions to the tens of nanometer dimensions, electron and other small wavelength beams must replace optical beams and correspondingly thin resists must be used. We describe this application in the next section.

Section III describes the use of Langmuir-Blodgett films as insulators in MISFET devices. The necessary requirement here is for a stable semiconductor-insulator interface with low interface density. We shall focus here on the results with InP as data on other semiconductors is most sparse.

Section IV describes the use of Langmuir-Blodgett films as tunneling barriers in Josephson junction electronic circuits. This application differs from the previous two in that it deals with a monolayer film rather than a multilayer film, and it concerns the transport of electrons and superconducting pairs through the film. Demands upon the film quality are extremely high in this application. This tunneling transport has more relevance to signal transmission in biological materials than do the other applications. And finally, the last section points out some of the many remaining questions in the applications of this developing technology.

II. LANGMUIR-BLODGETT FILMS AS ELECTRON RESISTS

Fabrication of integrated circuits with nanometer feature sizes requires the abandonment of photolithography due to its wavelength limitation and the use of x-ray, ion, and electron beams.[3] Resolution with electron beams appears limited due to the back-scattering and the range of energetic electrons which gives rise to a pear-shaped exposed volume of resist material. The pear dimensions are comparable to the range and thus to the desirable resist thickness.

For fifty nanometer feature sizes, it is desirable to reduce the electron beam range to this length by decreasing the beam voltage to the order of 2 kV and to use 50 nm ultra thin resists.[4] Conventional spin-coated resins appear to have an increasing number of pinholes for thickness less than 300 nm, thus limiting their usefulness.

Some multilayer Langmuir-Blodgett films have the requisite properties for use as ultra-thin resists. Namely, they appear homogeneous and of constant thickness, they appear pin-hole free, and they are plasma resistant. They also offer high sensitivity and good contrast.

As an example, 30 layer (90 nm) ω-tricosenoic acid films developed in ethanol have a sensitivity of 0.45 μC/cm^2 and a contrast of 3 parameters[5] which are competitive with that of PMMA. 60 nm grating lines on an aluminum substrate have been resolved with this resist. Plasma etching of a 150 nm thick aluminum film shows that 100 nm of ω-tricosenoic acid provides adequate protection. In reactive CCL_4 plasma etching, the relative etching rates of aluminum to tricosenoic acid is 6.7:1.[4]

Positive electron beam resists of multilayer Langmuir-Blodgett films of α-octadecylacrylic acid[6] have been reported with a feature size in the resist of 50 nm. Plasma resistance, sensitivity and constrast of these 20 layer (69 nm) films has not yet been determined.

The essence of these results is that polymerizable films can be deposited by the Langmuir-Blodgett technique which are of such high quality (pinhole free) that the possibility of ultra large scale integration (ULSI) with nanometer feature sizes appears promising.

III. MISFET

Metal-insulator-semiconductor (MIS) structures are important ones for field effect transistors (FET) and for photovoltaic diodes. A necessary requirement is that the semiconductor-insulator interface be electrically stable (and presumably, therefore, chemically and mechanically), and that the interface have a low surface state density. Otherwise, the electrical characteristics of such devices would drift with time and with applied voltages and the semiconductor surface potential would not be changed by the application of external voltages respectively. The success of silicon devices can to a large extent be attributed to the Si-SiO_2 interface properties.

There is a growing interest in the use of Langmuir-Blodgett films as passivating insulators on compound semiconductors[2] due to the fact that the native oxides, either thermally-grown or deposited, have not proved to be either stable or to have low interface state densities. Successful deposition of cadmium stearate onto InP[7] and onto InSb and (Hg-Cd)Te[8] narrow band gap semiconductors have been reported. FET device operation has been reported for the InP.

Au/Cr-$CdSt_2$-n-type InP[7] diodes with 30 insulating layers (78 nm) display conductance peaks which are dependent upon the method of substrate preparation prior to depositing the organic film. Hysteresis in the capacitance-voltage curves indicates charge trapping at the interface. The conductance and capacitance data yield a surface state density of 3×10^{11} states/cm^2-eV near the band gap center, a value which compares with 10^{10} states/cm^2-eV for the Si-SiO_2 interface. These results are quite promising considering the infancy of this technology.

The admittance-voltage characteristics of Au-diacetylene polymer-n-$Hg_{0.47}Cd_{0.53}Te$ MIS diodes with 31 monolayers show no hysteresis, a small negative flat band voltage, and a surface which can be accumulated, depleted, and inverted.[8] This Langmuir-Blodgett film transmits well in the far infrared region. The results offer promise for an MIS technology for narrow band gap semiconductors. The possibility of charge coupled device (CCD) image arrays operating in the infrared is most exciting.

IV. JOSEPHSON JUNCTION INTEGRATED CIRCUITS

Josephson junction integrated logic circuits depend upon the tunneling of electrons and of superconducting Cooper pairs between two superconducting electrodes separated by an insulator in the thickness range 2.5-5 nm. Larkins, et. al.,[9] have recently reported the observation of a dc Josephson current between lead alloy electrodes in which the insulator is a polymerized vinyl stearate monolayer deposited by the Langmuir-Blodgett technique. Subsequent observations have been reported[10] with Langmuir-Blodgett deposited polydiacetylene monolayers and with other electrode materials.

Langmuir-Blodgett tunneling barriers have the potential to provide the necessary control of the circuit parameters such that very large junction arrays will be feasible. A critical circuit parameter is the junction critical current I_o which, for the small junctions involved in such circuits, is the product of the critical current density J_c and the junction area. For supercomputer applications, the critical current I_o of the junctions in an array must be maintained within a 10% tolerance.

Unfortunately, the critical current density J_c is proportional to the tunneling probability of an electron through the insulating barrier, a probability that depends exponentially upon the barrier thickness and upon the square root of the tunneling potential barrier. This places very high control demands upon the barrier processing technology. One such technology that appears promising is the rf sputter-etching of lead alloy films in an oxygen plasma[11] in which precision control of the processing parameters is maintained.

The barrier thickness of Langmuir-Blodgett films is determined by the molecular weight of the chain-like monomers of the film or by the number of CH_2 groups in the chain. By careful control of the starting materials,[12] this number can be fixed although information on the variance of this number is unknown. Hence, by careful chemistry, this barrier thickness can be controlled. Moreover, by decreasing the number of CH_2 groups from that of the stearate films, the barrier thickness can be decreased and the critical current density increased which has attractive performance benefits. The lowest practical limits for the number of CH_2 groups is unknown at this time.

These polymerized Langmuir-Blodgett films appear to consist of large crystallites and to have a high degree of structural integrity.[12,13] This suggests that the tunneling barrier height is well controlled and is the same order of magnitude for these films in which the dominant volume is made up of CH_2 groups. The possibility of point defects such as vacancies exists but little information is available about them or their possible effects upon the tunneling probability.

Perhaps the largest uncertainty with these films is the possible thin oxide layer that can grow upon the base electrode prior to the Langmuir-Blodgett deposition. This layer would then be in series with the desired film, and the layer would decrease the tunneling probability and the critical current density. More experimental work is necessary on this point.

Two performance estimates are important in logic devices. One is the time required for switching. The second is the amount of energy involved in a logic operation. The time required is determined by the superconducting gap voltage of the electrodes, the dielectric constant and thickness of the insulator (junction capacitance per unit area) and the critical current density. The switching speed estimates for lead alloy electrodes with vinyl stearate barriers is 6 picoseconds.[10] Decreasing the number of CH_2 groups from that of the stearate film is expected to decrease this time estimate as the critical current density increases faster than the junction capacitance with decreasing barrier thickness.

The energy of a logic operation is the product of the device power and the switching speed. For lithographically-fabricated junctions, the device power level is expected to be of the order of 0.5 μwatt. This, when combined with the switching speed yields an energy of 3×10^{-18} joules. This is many orders of magnitude larger than thermal energy kT and six orders of magnitude smaller than the typical picojoule of semiconductor logic.

V. QUESTIONS

While the results to date in the application of Langmuir-Blodgett films to microelectronics have been most promising, many questions remain to be answered. One question of prime importance is the compatibility of these films and their depositions with prior and subsequent processing steps. It is questionable, for instance, whether such films can withstand even the low temperature processing steps associated with semiconductor integrated circuit fabrication. The use of polymeric films with strong intra-plane bonding may be key to their successful application. The same question must be answered relative to the compatibility with Josephson integrated circuit fabrication.

Electron resists studies have been carried out to ascertain the necessary material characteristics to obtain ultra high resolution. To date, however, device fabrication using these Langmuir-Blodgett resists has not been achieved. The potential of devices with nanometer features will undoubtedly push this area forward.

Perhaps the most exciting of the three application areas is the monolayer development for Josephson integrated circuits. The question of critical current control, while simply stated, has not been answered. The limit of film thinness is not yet known. Questions on the film thickness uniformity and on the defect density (in its effect upon tunneling) must be answered in more definitive detail than has been asked in the past. Perhaps the reason that this area is the most exciting is the implications such tunneling phenomena have for the electrodynamics (and statics) in biological materials.

ACKNOWLEDGEMENT

Some of the work reported here was done while the author was with Case Western Reserve University. The author acknowledges his many conversations on this topic with his colleagues, J. B. Lando, G. L. Larkins, Jr., C. W. Burkhart, and M. J. Deen. This research has been partially supported by the National Science Foundation through grant ECS81-20565.

REFERENCES

1. G. L. Gaines, Jr., Insoluble Monolayers at Liquid-Gas Interfaces, Interscience, New York, 1966.

2. P. S. Vincett and G. G. Roberts, Thin Solid Films 68, 135 (1980).

3. J. S. Greeneich, Electron-Beam Technology in Microelectronic Fabrication, p. 60, edited by G. R. Brewer, Academic Press, New York, 1980.

4. A. Barraud, C. Rosilio and A. Ruaudel-Teixier, J. Vac. Sci. Technol. 16, 2003 (1979).

5. A. Barraud, Thin Solid Films 99, 317 (1983).

6. G. Fariss, J. Lando, and S. Rickert, Thin Solid Films 99, 305 (1983).

7. G. G. Roberts, K. P. Pande, and W. A. Barlow, Solid State and Electron Devices 2, 169 (1978).

8. K. K. Kan, G. G. Roberts, and M. C. Petty, Thin Solid Films 99, 291 (1983).

9. G. L. Larkins, Jr., C. W. Burkhart, E. D. Thompson, J. B. Lando and E. Ortiz, Bull. Am. Phys. Soc. 27, 639 (1982); Thin Solid Films 99, 277 (1983).

10. G. L. Larkins, Jr., E. D. Thompson, M. J. Deen, C. W. Burkhart, J. B. Lando, IEEE Trans. Magn. MAG-19, 980 (1983); Bull. Am. Phys. Soc. 28, 568 (1983).

11. J. M. Greiner, et. al., IBM J. Res. and Devel. 24, 195 (1980).

12. A. Cemel, T. Fort, Jr., J. B. Lando, J. Polym. Sci., Part A-1, 10, 2061 (1972).

13. D. R. Day and J. B. Lando, Macromolecules 13, 1478, 1484 (1980).

NONLINEAR TUNNELING BARRIERS AT HIGH FREQUENCIES AND THEIR POSSIBLE LOGIC PROCESSING FUNCTION IN BIOLOGICAL MEMBRANE

Gerald C. Huth

University of Southern California
Institute for Physics and Imaging Science
4676 Admiralty Way #932
Marina del Rey, CA 90291

James D. Bond

Science Applications Inc.
1710 Goodridge Dr., P.O. Box 1301
McLean, VA 22102

Per Arne Tove

Uppsala University
Teknikum, Box 534
751 21 Uppsala, Sweden

INTRODUCTION

One way to visualize the complex phospholipid molecular bilayer comprising the biological membrane which surrounds the living cell is as a high density, two dimensional, digital logic processing (possibly memory) plane or surface acting somewhat in the fashion of an "ultra-microcircuit" composed of high speed, low power-consuming, quantum-limited units ("switching elements" or "devices") which would be composed of a bilayer "sandwich" of two opposed, single linear phospholipid molecules. In this view, the analogue of silicon in traditional microcircuitry would be the inner insulating hydrocarbon acyl chain layer (approximating 40-50 Angstrom units in thickness). Surface protein molecules or the conducting surface layers of assemblies of the polar head groups of the amphipathic membrane-forming molecules would form a sort of "conductor/insulating barrier/conductor" structure already familiar as MIM, MOS and other more complex types of metal/semiconductor devices. Such a thin structure with the high electric field (10^7 v/M) characteristically present in living cell membrane presents the possibility of electron tunneling between the conductive regions. A lateral extension of membrane surface proteins or, in another hypothesis, the dipole moments of the polar head groups of the phospholipid molecules might be viewed as comprising the necessary "architecture" (again in microcircuitry terms) for a logic processing circuit. Of course, living systems may have evolved some sort of architecture more advanced than our current microcircuitry state-of-art. Indeed the Very Large Scale Integrated Circuitry (VLSI) developmental effort is currently searching for improved logic addressing and processing architectural concepts. Of one additional note - biological membrane would appear to possess a "double-sided" architecture - not at all used in current semiconductor technology.

The spirit of this paper is not to argue necessarily that nature is behaving in this fashion, i.e., using on-off molecular "switches" to perform standard Boolean operations. The computational aspects of biological systems are probably a great deal more complex than that. For example, such computation probably takes place in three dimensions in biological systems - using enzyme/substrate reactions as an example (termed by the biologist "recognition"). A comparison between two dimensional semiconductor logic and the biological membrane structure, however, can be made and it may provide some insight as to a possible "biological computer".

The aspect of the dynamics of living systems must be introduced to provide the basic for the hypothetical necessity for "computation". Prigogine (I. Prigogine et al., 1977) and Frohlich (H. Frohlich, 1980) have developed hypotheses from the most fundamental point of view regarding the possible mechanisms that such systems utilize in creating order through the dissipation of metabolic energy. Del Giudice (E. Del Giudice et al., 1982) proposes that the cyclical behavior characteristic of living systems involves two mechanisms the first being the aforementioned one followed by longer term storage of energy on biopolymers in the form of solitons as proposed by Davydov (Davydov, A. S., 1980). An aspect of Frohlich's theory is that high frequency phonon-like vibrations in the region of 10^{10} - 10^{11} Hz are present in the membrane of living cells and other biological macromolecules. Microwave resonance effects at similar frequencies in living cellular systems have been reported (Devyatkov et al., 1974, Grundler et al., 1983) and further implicate high frequency oscillations in biological processes. An attractive line of reasoning proceeeds from the presence of this high frequency oscillatory behavior in living systems. Its' possible use in a logic processing system with perhaps unique properties forms the basis of the ideas presented in this paper.

Some additional thoughts regarding the state-of-art of logic processing technology are presented in Table I. Digital semiconductor (primarily silicon) logic is entering the sub-micron, minimum feature size. In contrast, a molecular or biological system might make molecular diameter (of the order of square Angstrom units) elements possible. Of great importance, in any such technology is the voltage excursion required across whatever the intrinsic, switching nonlinearity might be. In silicon this value is 600 millivolts. Cryogenic "Josephson" logic currently under development for computer application requires typically only 3 millivolts to effect this transition - and this coupled with speed is its' primary advantage. Biologic logic based on room temperature electron tunneling (or even soliton transport) would have to require values of this order or less. Speed of response of all logic systems is related to the voltage excursion CV/i where C is the capacitance between output terminals and i is the current available to drive the change of state. Both Josephson and hypothetical biological logic would be extremely fast - in the 10^{-12} sec (or picosecond) time domain.

RELEVANT ASPECTS OF BIOLOGICAL MEMBRANE STRUCTURE AND CONCEPTUAL ARCHITECTURES OF A LOGIC PROCESSING SYSTEM

The most ubiquitous constituent of the membrane surrounding cells is a two dimensional matrix of phospholipid molecules arranged in a bilayer configuration. Its thickness is of the order of 50-80 Angstrom units. On the surface of, and embedded in, this bilayer are protein molecules of a wide variety of chemical composition and geometrical conformation. Proteins on the surface of the membrane are referred to as "peripheral" proteins. Embedded protein molecules

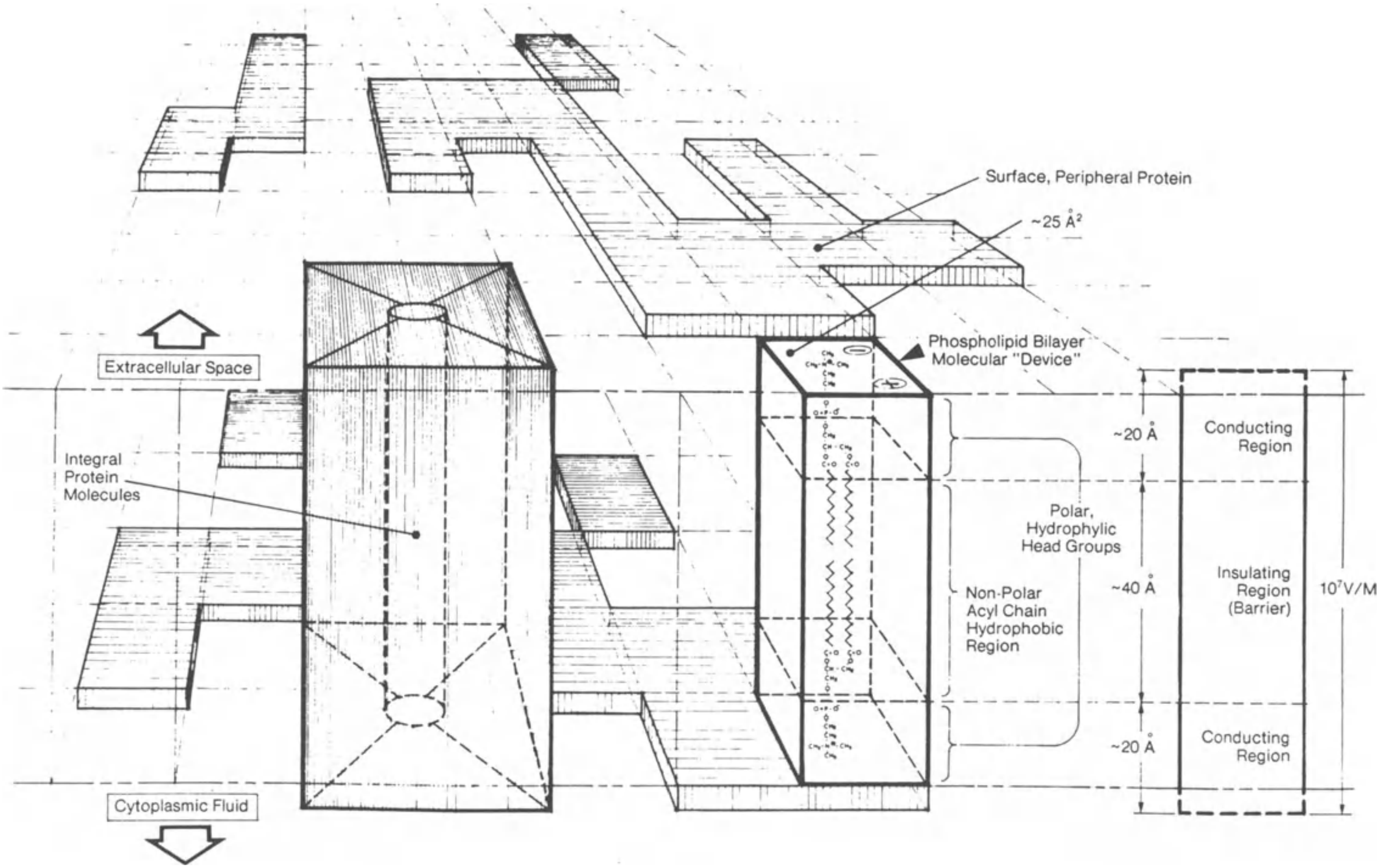

Surface, Peripheral Protein
~25 Å²
Phospholipid Bilayer Molecular "Device"
Extracellular Space
Integral Protein Molecules
~20 Å
Conducting Region
Polar, Hydrophylic Head Groups
Non-Polar Acyl Chain Hydrophobic Region
~40 Å
Insulating Region (Barrier)
10^7V/M
~20 Å
Conducting Region
Cytoplasmic Fluid

Figure 1. Generalized, highly schematic concept of biological membrane structure

which can span the entire thickness of the bilayer are referred to as "integral" proteins. In most biological membrane systems there is an asymmetric distribution of lipids in the two monolayers that comprise the bilayer. Each monolayer usually contains the same kinds of lipid molecules but different amounts of each. With regard to the structure of individual phospholipid molecules, they are composed of a phosphatidic "polar head group" - the hydrophilic segment of the molecule- and two nonpolar, hydrocarbon "acyl" chain groups - the hydrophobic portion. The latter give the membrane its' extraordinary insulating properties.

The polar head group of these molecules may be either negatively charged or zwitterionic (dipolar) at physiological pH. In a generiic sense "polar head group" refers to either of the two cases. It is also known that the acyl chains can exist in different isomeric states.

Regarding the structure of the proteins associated with membrane, they too are amphipathic. They, however, can possess much higher (an order of magnitude) dipole moments than their lipid counterparts. In one hypothesis (Baumann, 1983) the amphipathic nature of membrane spanning proteins is thought to be utilized laterally, i.e., in the plane of the membrane, as a "controllable switch" to gate ion flow related to the metabolic energy transduction process.

Related to potential fabrication of molecular devices, it is possible to fabricate a "fascimile" of at least the basic (core?) phospholipid structure of biological membrane using the "self-organizing" character of such molecules by the BLM technique (Tien, 1974). It then becomes possible to study bilayers that structurally approximate biological membrane. This method currently seems limited to low frequency "ionic conductance mediated" measurement because of the water medium in which these films are grown.

The concept of electron tunneling in biological systems seems to have been experimentally verified in a number of instances. De Vault (De Vault and Chance, 1966) measured an unusual temperature independence (to $T < 130^{o}$ K) of the velocity of electron transfer from Cytochrome to the Chromatin reaction center. Such temperature independence suggests a tunneling mechanism. Lewis (Lewis, 1982) using BLM films of phosphatidyl choline has measured voltage -

current characteristics that he interprets as electron tunneling. Phosphatidyl choline or lechithin is a component of the membrane of living cells. Electron tunneling has also been reported in thin, ordered Langmuir-Blodgett films of polymerized vinyl stearate (Larkins, et al., 1983) although these measurements were made at cryogenic temperatures. The Langmuir-Blodgett technique is similar to the BLM technique in that both take advantage of the self-ordering nature of amphipathic molecules in or on the surface of water. All of these results seem to indicate that electron tunneling can occur along organic, carbon based molecules in structures approximating biological membrane.

The concept of how a two dimension array of molecular tunnel elements might resemble logic processing plane is shown in Figure 1 in a schematic way that emphasizes the electrical nature of membrane. A basic switching unit or electron tunnel "device" might be the single, phosopholipid bilayer element composed of the two amphipathic molecules. (The chemical structure of one such representative molecule - phosphatidyl choline - is superimposed). The area encompassed by such a device would be about 25 square Angstrom units and in excess of 10^9 of them would, in two dimensional array fashion, form the membrane of a typical living cell. The device and general membrane core structure appears in cross section, as indicated, to be similar to a conductor/insulating barrier/conductor "sandwich" alluded to earlier. Biological membrane displays quite extraordinary electrical properties as pointed out by Frohlich (Frohlich, 1975). The nanometer dimension and very high electric field present conceptually support the possibility of electron tunneling.

As noted earlier and indicated in the figure, protein molecules span portions of both surfaces and in some locations pass entirely through the membrane layer. We have taken great latitude in drawing the shape of the peripheral proteins but the point is: might the contact area and resultant shape of these proteins represent a lateral "architecture" electronically contacting groups of electron tunnel devices? Alternatively, large assembledges of groups of protein molecules may form a greater architecture. Some of the integral transmembrane proteins are undoubtedly involved in the Na-K ion "pumping" system that provides metabolically derived energy to "drive" the entire system. In an electrical sense such molecules might

Table I

	SEMICONDUCTOR-(SILICON) MICROCIRCUITRY LOGIC	JOSEPHSON SUPERCONDUCTING LOGIC	BIOLOGICAL LOGIC
MINIMUM FEATURE SIZE	0.5 μm(5,000 Å) (MOS)	?	25 Å(?)
VOLTAGE CHANGE ACROSS NONLINEARITY	600 mv	3 mv	?
SPEED (CV/i)	10^{-10} - 10^{-11} sec	10^{-11} - 10^{-12} sec	10^{-12} sec
POWER DISSIPATION	1	$\frac{1}{250 - 5,000}$	$\ll \frac{1}{250 - 5,000}$

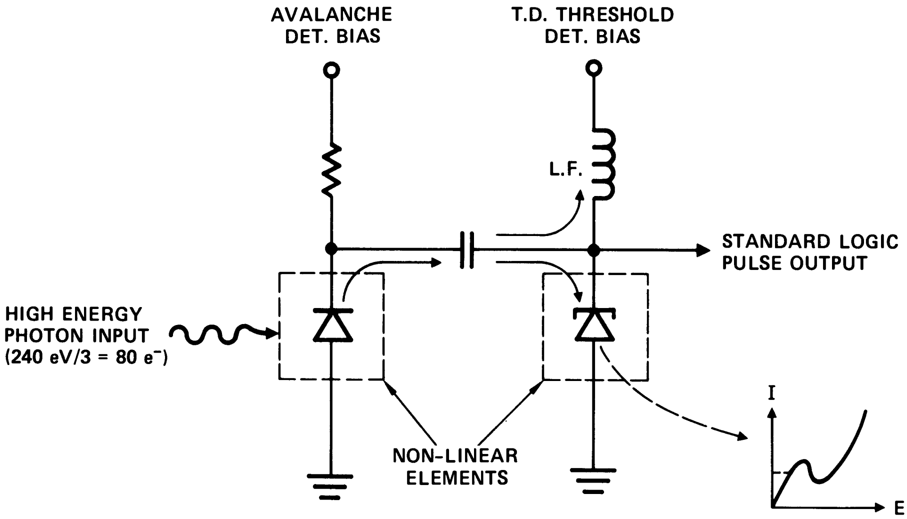

AVALANCHE DET. BIAS
T.D. THRESHOLD DET. BIAS
L.F.
STANDARD LOGIC PULSE OUTPUT
HIGH ENERGY PHOTON INPUT (240 eV/3 = 80 e⁻)
NON-LINEAR ELEMENTS
I
E

Figure 2. Illustration of use of fast time domain to descriminate against noise.

appear as conductive "feedthroughs" and are even reported to possess a degree of spatial (hexagonal) order.

Another possible concept that might lead to an architecture of sorts has been proposed (Bowen and Lewis, 1982) in which the dipole moments of the polar head groups of the phospholipid molecules - which can be oriented somewhat laterally toward the plane of the membrane - might form domain structures in that plane. These dipole moments can be quite large - in the vicinity of 20 Debye units.

If biological membrane were behaving as a logic processing system it would, judging from the size of its structural elements and the speed characteristic of electron tunneling phenomenon, be operating in the high frequency, very fast time domain. A logic processing system operating in this region, specifically the picosecond (10^{-12} second) domain, would possess unusual attributes that the living system would conceptually have found it necessary to evolve. It might be, for example, to a large degree immune to influences such as operation at elevated temperature and visible photon interaction effects which would interfere with a slower, lower frequency system. One of the authors (Huth) described (Locker and Huth, 1966) an ionizing radiation detection circuit that may have provided a glimpse of this type of unusual behavior. The basic circuit described in that paper is shown in Figure 2. In that circuit, a subnanosecond (less than 10^{-9} second) ionizing radiation/thermal electron transducer (a high speed "avalanche" type of amplifying solid state radiation detector) is capacitively coupled to an equally fast, silicon or gallium arsenide tunnel diode element. The overall circuit then becomes a transducer from an initial picosecond (10^{-12} second) event (the stopping, time or charge generation time for an ionizing interaction in a solid) to a standardized, 1-2 volt logic output in only two circuit elements. It then becomes in a sense the high speed analogue of the Geiger Mueller counter. The unusual behavior as noted in the 1966 paper and in subsequent measurements, was the ability to count events involving extremely small numbers of electrons (for example, a pulse of 80 electrons from a 242 ev carbon x-ray interaction) with the detector leg of the circuit at temperatures elevated above room temperature. Alternatively, such detection could be accomplished with flux of rather intense visible light incident on the detector again creating a very high DC leakage current through the detector leg. It is of note that detection of ionizing radiation events of this small magnitude generally requires cryogenic cooling to reduce noise. In this circuit, not only was it true that cooling was not required, but the detector could be actually heated while maintaining detection of these extremely small events.

Such behavior we believed to be explainable in that one form of noise (termed of the "parallel" type in nuclear electronics terminology) measured in this nanosecond time constant possesses a "time integratable" property being reduced as time constant is shortened. What becomes important then in the fast time domain in which this circuit operates is the number of single noise electrons competing with the event to be detected in the short "detection logic" time of the system. Thermally induced current then begins to become bothersome when the number of electrons per detection-logic time become large enough to begin to statistically interfere with signal. An optimized system for nuclear ionizing radiation detection should ideally possess a time constant of approximately 10^{-12} second if such could be accomplished. Such a system would make events corresponding to even the softest x-rays most unique in nature and detectible under a variety of outside influences.

It should be noted that, in contrast to "parallel" noise which is reduced as a time window is shortened, another form of noise (termed "series" noise again in nuclear electronics parlance) exists which has the characteristic of increasing in magnitude as time constant is shortened. This type of noise generally dominates at our current technological state in the sub-nanosecond time domain. It is physically occasioned, for example, by the finite conduction length of the base of a field-effect transistor. The general effect in the vicinity of any slower element in a logic-amplifying chain chain, therefore, is for events to "pile up" and thus increase in magnitude as measured in some slower time frame. A biological membrane may have overcome this problem with its' exquisitely close packed, high frequency optimized structure.

It is, of course, obvious that no matter to what degree the effect of noise events can be minimized by operating in the fast time domain the concept would seem to break down at the single electron level when a noise "event" becomes indistinguishable from a "signal electron". The only solution to this in the biological concept might be that "preamplifying,Frohlich-type front-ends" would serve to amplify an input signal above the noise level (in much the same manner as the avalanche detector did in Figure 2). Another possibility might be that some directional aspect may be operative in the electron tunneling process (such as in the photon/electron conversion process described in the next Section) to increase the information bearing content of this phenomenon.

ROOM TEMPERATURE ELECTRON TUNNELING THROUGH BARRIERS AT HIGH FREQUENCIES IN SEMICONDUCTOR JUNCTIONS AND ITS POTENTIAL RELEVANCE TO BIOLOGIC LOGIC

A fundamental question in considering the possibility of biological logic processing is: Can the information transferring entity - the electron in our hypothesis - be conceptually "isolated" or defined so as to be able to perform discrete device functions (detection, amplification, switching, memory). Concurrently it is required that such isolation be accomplished at the operating temperature of the system - in this case slightly above room temperature.

Single electron definition or measurement at room temperature has been accomplished to this time in (to our knowledge) only one instance. This uses the unique combination of the high gain and low noise and fixed geometrical aspect of the photomultiplier tube. In this device, the single electron leaving the photosurface is amplified by the very high gain of the device (10^7) with the result being a unique, identifiable output pulse which corresponds to the event. (Since the photon/electron efficiency of the photosurface is far less than unity, however, it cannot be said that this corresponds to a single photon/electron conversion.)

In an analysis which projects the performance of inorganic semiconductor (Schottky type) nonlinear tunneling barrier diodes to high frequencies Tucker (Tucker and Millea, 1978) shows that such devices possess sufficient nonlinearity at room temperature so as to display quantum limited behavior at wavelengths from the visible extending into the infrared - they project to 25 microns. These projections we believe to have relevance in two areas relative to the possibility of biological logic. First they indicate that it is possible to attain sufficiently nonlinear behavior at room temperature to attain the quantum limited state - the "isolated" or defined electron discussed previously. Interestingly, high frequency is required. Secondly, Tucker focuses on photon-electron conversion which may have relevance to the microwave cellular interaction effects described earlier. In this regard, biological systems will require a greater degree of nonlinearity than displayed by semiconductor Schottky barriers.

Schottky barriers possess exponential voltage-current characteristics:

$$I_{dc}(V) = I_0 \exp(SV) \qquad (1.)$$

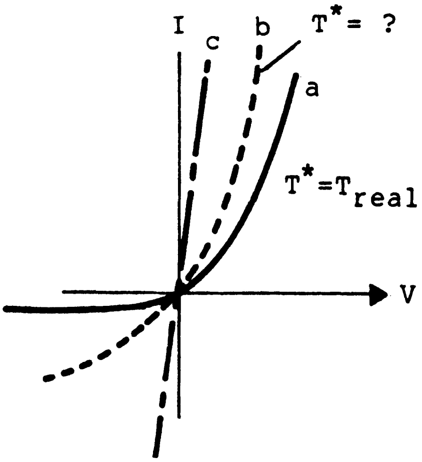
where the coefficient of nonlinearity S is expressed by:

$$S = e/kT \qquad (2.)$$

In high frequency applications (of the type we are interested in), heavy doping is required to produce a degenerate barrier sufficiently thin so that electron tunneling near the Fermi surface becomes the dominant current carrying mode. S becomes practically independent of temperature in this case and can be expressed by:

$$S \equiv e/kT^* = 11600/T^*(^\circ K)V^{-1} \qquad (3.)$$

Equation (1.) represents the forward characteristic of a Schottky barrier when $eV \gtrsim 3kT/e$. At room temperature this results in $S \lesssim 40V^{-1}$. The corresponding voltage-current characteristic is shown as (a):



If the doping is made so high that thermally assisted tunneling or even pure tunneling occurs, the characteristic changes from (b.) to (c.) reflected in changes in the value of S.

In normal semiconductor diodes the slope of the IV curve is governed by the energy bandgap, of the order of 1 eV. For a superconducting diode the "bandgap" is of the order of kT_C which is of the order of a few millivolts. The nonlinearity thus is much sharper and one may arrive at values of S of 10^4 (which extends room temperature quantum limited response to longer wavelengths).

Tucker derives an expression for the "current responsivity" of a tunneling barrier defined as the change in DC current per unit power absorbed as:

$$R_i^{QM} = \frac{e}{2kT^*} \frac{\tanh(\hbar\omega/2kT^*)}{(\hbar\omega/2kT^*)} \qquad (4.)$$

This reduces to the classical expression:

$$R_i^{CL} = \frac{e}{2kT^*} \qquad (5.)$$

when the conductance varies slowly on the voltage scale $\hbar\omega/e$. At high frequencies, however, where conductance does vary rapidly the quantum current responsivity approaches the fundamental limiting value. Tunneling barriers then are energy detectors at low frequencies. At high frequencies, however, the quantum limit is reached $\hbar\omega \gtrsim 2kT^*$ "where the net absorption of a single photon provides sufficient energy to allow one additional electron to tunnel in the preferred direction" (a quote from Tucker).

A coefficient of nonlinearity S of 40 V^{-1} (or 20 V^{-1} which can be obtained for a gallium arsenide diode because of a strong contribution from recombination current) is sufficient to reach quantum limited responsivity to a wavelength of 10 microns. S values of 2000-4000 would be required to extend this behavior to the millimeter wavelength region. As noted above, superconducting tunnel devices display S values of 10,000. These values seem to set the bounds on what characteristics a biological "device" will have to display.

CONCLUSIONS

We have presented a certainly highly speculative argument as to how electron tunneling which has been reported in biological molecules might be utilized by living systems in some sort of organized, higher order, logic processing fashion perhaps somewhat as is understood in the current technology of semiconductor microcircuitry. If some sort of logic processing were operative, we propose that the switching entity may be the electron although a number of other groups are now speculating that the possibly experimentally verified concept of the "soliton" may accomplish the same function. Suffice it to say that whatever the entity turns out to be it must be of exceedingly low mass and require a minute amount of energy to perform the switching function. Regarding biological membranes and living cellular systems, the work of Frohlich in introducing high frequency and quantum electrodynamic theory to this area is certainly of fundamental importance. Experimental evidence of resonant microwave interactions with living cells further implicates high frequency in these processes.

The fundamental question as to whether or not a system for the transfer and processing of information need exist in living systems is indeed a crucial one. Some argue that the idea is not necessary at all and that genetic, chemical and the self-replicating nature of biopolymers will be sufficient to eventually explain biological action. We believe that perhaps effort at devising very clever experiments to measure the general presence of high speed electronic processes in living cellular systems may be the best near term approach to arriving at an answer to these questions.

We might finally speculate that unless Frohlich is correct about high frequency oscillatory behavior being present in living biological systems, we will not arrive at the logic described in this paper and the "super-biological computer" envisioned and discussed of late by many.

REFERENCES

Baumann, G., (1983) "Molecular Switching in Neuronal Membranes" 2nd International Workshop on 'Molecular' Electronic Devices", Naval Research Laboratory, Wash. D.D., April 1983.

Bowen, P. J. and Lewis, T. J., (1982) "Electrical Interactions in Phospholipid Layers," First International Conference on Langmuir-Blodgett Films, Durham, Great Britain, September 1982.

Davyatkov, N. D., (1974) Sov. Phys. Ups. 16, 568, (1974)

Del Giudice, E., Doglia S. and Milani M., "A Collective Dynamics In Metabolically Active Cells" Physica Scripta, 26, 232-238, (1982).

De Vault, J. D., Chance, G. O., (1966) Biophys. Jour., 6, p. 825.

Frohlich, H., (1975) "The Extraordinary Dielectric Properties of Biological Materials and the Action of Enzymes" Proc. Nat. Acad. Sci, USA, 72, No. 11, pp 4211-4215, November 1975.

Frohlich, H., (1980) "Advances in Electronics and Electron Physics" Ed. L. Marton, C. Marton, 53, 85, (1980).

Grundler, W., (1983) "Nonthermal Resonant Effects of 42 GHz Microwaves on the Growth of Yeast Cultures" Coherent Excitation in Biological Systems, Eds. H. Frohlich and F. Kremer. pp. 25-46, Springer Heidelberg, 1983.

Lewis, T. J., (1982) "Electronic Processes in Biology" Phys. Med. Biol. 27, pp 335-352, 1982.

Larkins, G.L., Thompson, E.D., Ortiz, E., Burkhart, C.W. and Lando, J.B., (1982) "Langmuir-Blodgett Films as Barrier Layers in Josephson Tunnel Junctions" First International Conference on Langmuir-Blodgett Films, Durham, Great Britain, September, 1982

Locker, R.A. and Huth, G.C., (1966) "A New Ionizing Radiation Detection Concept Which Employs Semiconductor Avalanch Amplification and the Tunnel Diode Element" App. Phys. Letters, 9, September 5, 1966.

Prigogine, I and Nicolis, G., (1977) "Self Organization in Non-Equilibrium Systems From Dissipative Structures to Order Through Fluctuations" Wiley & Sons, New York, New York, (1977).

Tucker, J.R., Millea, M.F. (1978) "Photon Detection In Nonlinear Tunneling Devices" App. Phys. Lett. 33, No. 7, 611-613, October 1978.

MOLECULAR ELECTRONICS: AN OPPORTUNITY FOR A BIOTECHNICAL SYNERGISM

Forrest L. Carter

Chemistry Department, Code 6175
Naval Research Laboratory
Washington, D.C. 20375

I. INTRODUCTION

The scientific, technological, and humanitarian opportunities inherent in the strong interaction between molecular biology (and biochemistry) and the developing new field of molecular electronics are unusually abundant and stimulating. It is an object of this paper to discuss a sampling of these opportunities and to indicate their importance, which in some cases, are revolutionary in scope.

The continuing success of the semiconducting industry in ever reducing the size of their switching components (i.e., transistors) suggests that in a few short decades such switching elements will be at the molecular size level [1]. Although A. Aviram and M. Ratner discussed a ''molecular rectifier'' in 1974 [2], it wasn't until 1979 that the concepts, prospects and problems of Molecular Electronic Devices (MEDs) were outlined [1a]. This has been followed by two MED workshops which were internationally attended [3]. The contribution of molecular biology and related fields to the conceptual development of molecular electronic devices has been rewarding, as this paper will show, and these contributions will undoubtedly continue to enrich the field.

Areas of probable interaction between biotechnology and molecular electronics are listed in Table 1. While some of these topics like item 1 (Merrifield synthesis), item 4 (visual data processing), and item 5 (molecular lithography) will be the subject of individual sections below, others will be discussed less formally when the context is appropriate. Each of the

TABLE 1

1. Application of Merrifield-Type (Modular) Synthesis to MED Fabrication.

2. Learn from Molecular Biology Principles of -

    A. Self-Organization
    B. Self-Synthesis
    C. Biological Switching
    D. Energy and Light Harvesting

    and apply to organic and inorganic synthesis.

3. Biotechnical Synthesis of Chemical Subunits for later MED assembly.

4. Parallel Processing of Visual Data.

5. The use of Designed Biological subunits for Molecular Lithography.

6. Attachments of Monoclonal Antigen and Enzyme Fragments in a Hybrid Semiconductor-MED detector.

7. Molecular Biology and MED Technology combined in Prosthetic Devices.

8. Digital to Content Addressable Memories.

9. Reliable Control Systems and Neural Models.

topics of Table 1, however, represents a very promising synergistic technological and scientific opportunity. In my mind, however, the most important is the second item.

If it is possible to learn the principles of self-organization and self-synthesis from the biological world and apply those principles to organic and inorganic chemistry then we will be opening a door on a new and powerful technology. Not only will our scientific knowledge be immeasurably enhanced but it would make possible a synthetic and microstructural fabrication technology capability of even greater importance than the successful development of an MED technology. In short it would permit the design and fabrication of materials whose optical, electrical, structural, etc., properties which can be varied over nanometer dimensions. While no further consideration of the synthetic aspects of second item will be made in this paper, it is the underlying assumption of a MED technology that the challenge offered in this paragraph is at least partially met.

The fabrication needs of molecular electronics are briefly outlined below in section II where the proposed synthesis of molecular wires of (SN)x are given as an example of a ''modular'' approach (Table 1, item 1). A discussion of MED switching concepts follow: electron tunnel switches are treated in section III and conformational switching, especially soliton switching, is discussed in section IV. The opportunities possible in the use of soliton valves and cellular automata are discussed in section V while section VI considers soliton generation in relationship to parallel processing of visual data (Table 1, item 4). Opportunities for resolution at the molecular size level in lithographic techniques and their relation to biotechnology is then considered in section VII. This area of molecular structuring is further treated in section VIII using ''organizing'' molecules. Finally, section IX considers items 6 to 9 of Table 1 as other areas where synergistic opportunities are certain to occur.

II. MED FABRICATION: A MERRIFIELD-TYPE ''MODULAR'' SYNTHETIC APPROACH

MEDs seemed technically feasible only after it was recognized by the author [1a] that their preparation might be feasible using an approach based on the Merrifield synthesis of polypeptides. While other techniques of MED fabrication have been considered [4,5] only the ''modular'' or Merrifield type synthesis will be considered here.

The ''modular'' preparation of a MED complex is projected as follows: On a very clean substrate, lithographic techniques are used to define different areas for the attachment of conducting molecular ''wires'' and for insulating layers. Then using the techniques of surface modification chemistry [1a] the first subunits of the molecular ''wire'' are bonded covalently to the desired conducting areas of the substrate. In a different series of chemical reactions the insulating areas are prepared and then reacted with the insulating monomer to form an initial layer. For both types of materials the substrate is added to by using reagents that add only a single subunit at a time, hence the designation ''modular'' chemistry. The preparation of conducting filaments of (SN)x might be achievable by alternately applying sulfur diimide $((CH_3)_3SiN)_2S$ and then SCl_2 to the substrate [6]. This approach derives from the observation of J. Milliken that oligomers of (SN)x can be prepared from the diimide and sulfur dichloride [7]. The modular approach reacts each reagent separately with the reactive end of the filament thereby adding to the length of the filament while preparing the filament end for the next reagent. By repeating the reagent addition sequence n times one increases the chain length of (SN) by exactly $n(SN)_2$ units. By this technique the exact length of the (SN)x chain can be specified by the number of reactions. Fig. 1 illustrates the production of molecular wire attached to a conducting silicon substrate. Similarly molecular wires of trans-polyacetylene (CH)x can presumably be formed by a series of two different ''modular'' condensation reactions after the first unit has been covalently bonded to clean Pt metal by surface modification chemical techniques.

Once the substrate had been covered with an initial set of molecular electrical connections and insulation then another series of reactions would be employed to add the memory and switching components. Thus, layer by layer different components would be added to different areas by a series of chemical reactions. Of course to make a three-dimensional functional architecture, the new added components must lie down on the growing substrate in the correct conformation and then be stitched into place to form a solid structure. The differentiation of the substrate surface implied in the above will be considered in section VII.

Upon adding a chargable chromophore at the terminus of an -(SN)n-molecular wire bonded to a conducting silicon substrate one has a testible connection between the macroscopic and the molecular world. It is testible because in the presence of an electric field or in an electrochemical cell the chromophore could pick up a charge with an associated detectable change in its absorption spectra. At this point one has no obvious means of molecular addressability. However, addressability can be achieved by the insertion of switches. This aspect of MED technnology will be considered in the next three sections.

III. MOLECULAR ANALOGUES OF NAND AND NOR GATES

Electron tunnelling through a short series of periodic barriers has been proposed as the basis for molecular analogues to both semiconductor NAND and NOR gates [1,8,9]. This possibility was based on the earlier theoretical treatment of electron tunnelling by Pschenichov [10]. He showed in a quasi-classical argument, that if the energy of an electron approaching a short periodic array of barriers matched exactly the energy of one of the resonant or pseudostationary states, that the electron would penetrate the barriers with a perfect transmission coefficient (Fig. 2). In 1981 Pschenichov's argument was supported by an exact quantum mechanical treatment using potential step functions [9] and the switching model was confirmed recently numerically [11-13].

In Fig. 2a an electron is shown approaching a short array of periodic barriers from the left. If the energy of the electron

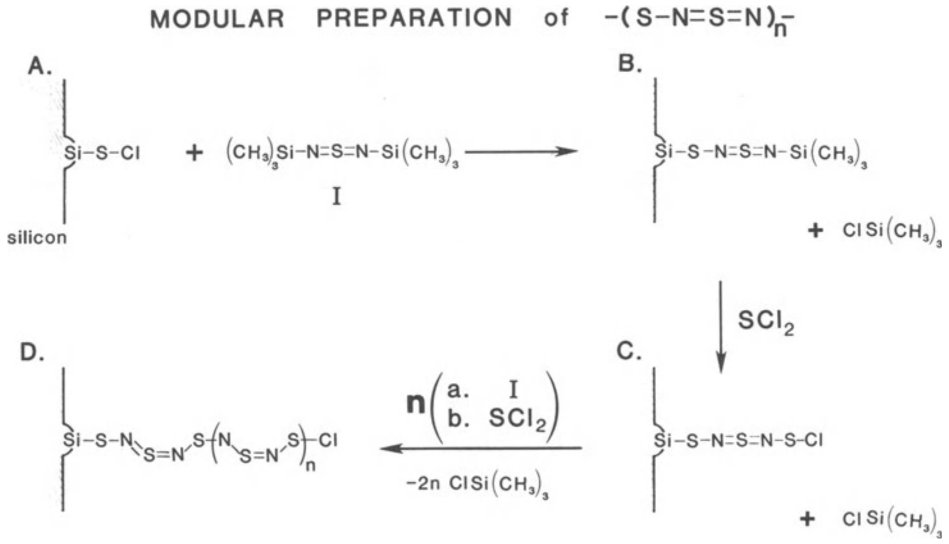


Fig. 1 In this proposed synthesis the length of molecular wires of $(SN)_x$ can be controlled exactly, in this case, to n+1 $(SN)_2$ units.

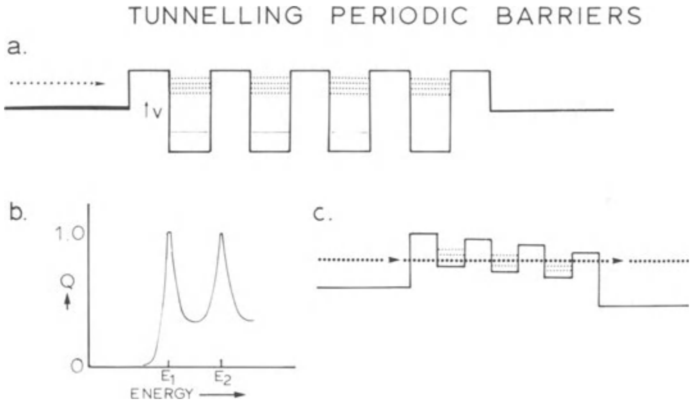


Fig. 2 Resonant tunnelling via E_i pseudo-stationary energy levels is indicated in (a) and (c). The transmission Q(E) decline as the electron energy deviates from E_i is schematically indicated in (b).

matches the pseudo-stationary state energy levels (dotted lines between the barriers in Fig. 2a and E_1 and E_2 in Fig. 2b) then transmission coefficient Q goes sharply to 1.0 (perfect transmission) as in schematic diagram of Fig. 2b. Excellent transmission can also be obtained in the case of Fig. 2c if the potential drop, from one well to the next, matches the resonant level energy differences [13].

The molecular analogue of the semiconducting NOR gate (Fig. 3b) is shown (Fig. 3a) as a fluorine-bridged gallium compound similar to that reported for phthalocyanine by Kuznesof, Wynne, Nohr, and Kenney [14]. The potential wells are defined by the positive charges of the nearby quaternary nitrogens while the barriers correspond to the fluorine bridges. These quaternary nitrogens are considered to be Control Groups (CGs) most of which are permanently fixed and considered to be dummy Control Groups. These CGs are mounted on rings as in Fig. 3c with either -(SN)x- leads (Control Groups) or with $-CH_2SO_3-$ groups (dummy Control Groups).

The D rings (Fig. 3d), nickel phthalocyanine moieties, provide links to both the ground and the negative potential through Ni-S bonds (Fig. 3a); as well as to the -(SN)n- Output lead. These terminating rings also serve to provide a suitable environment for the stacked bridged phthalocyanine rings. Control through the Input -(SN)n- leads A or B occurs via an electron flow down the -(SN)n- chain to neutralize the quaternary imino nitrogen. The neutralization of either nitrogen would change the potential of that ring sufficiently to sharply cut off the electron tunnelling through the rings.

Although the variety of possible control groups is large and has been noted elsewhere [1b] it will be useful to illustrate with the example of Fig. 4. Here we see the Control Group ring has two cyanine-dye like CGs, the one on the right with the quaternary nitrogen charge localized on the cyanine dye nitrogen furthest from the ring. In the context of a tunnel switch as in Fig. 3a this would turn the electron tunnelling off, where as the left CG of Fig. 4 is in the ''on'' position. Note that by having the different lengths of conjugation for the left and right CGs the two CGs could be switched by photons of different wavelengths. Accordingly, the incorporation of the ring of Fig. 4 with a system like that of Fig. 3a provides one with a photoactivated NOR gate. In the next section we will indicate an approach to the control of charge location in cyanine dye type Control Groups via soliton propagation. Counter anionic charges for the ring system might be built into the R groups.

The obvious chemical complexity of the structures indicated in Figs. 3 and 4 are but a small measure of the overall synthetic challenge inherent in the production of MEDs. One future solution to this synthetic problem may be that suggested by K.M. Ulmer [5,15, see also Ref. 16] of using a biotechnical preparation of chemical subunits for later MED assembly (Item 3, Table 1). While even self-

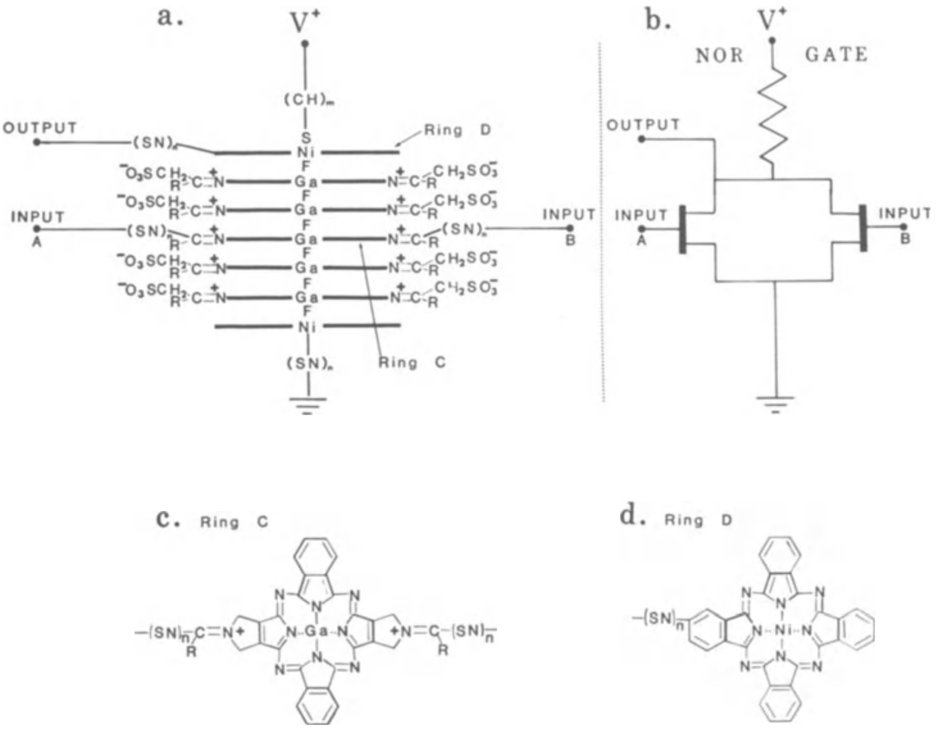


Fig. 3 A molecular analogue of a NOR gate (b) based on a stacked fluorine bridged ring system.

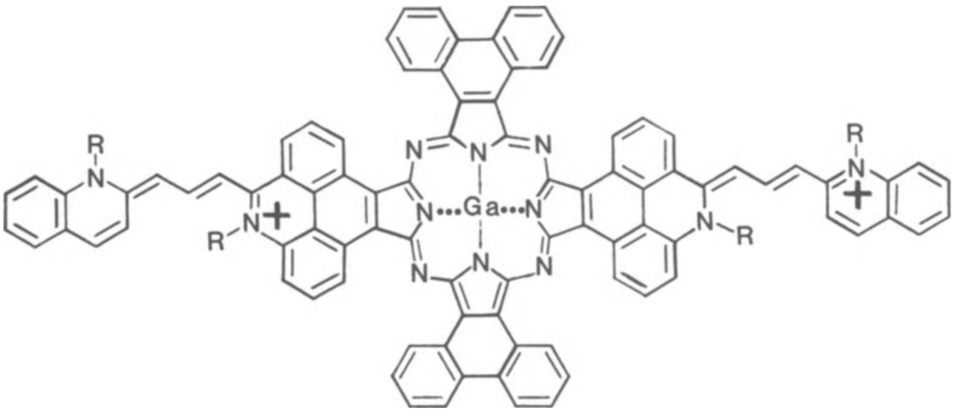


Fig. 4 Cyanine-dye like Control Groups on a ring of an electron tunnel switch.

assembling components have been discussed it should be remembered that currently recombinant DNA synthetic techniques are only applicable to the preparation of proteins. That is, the DNA coding necessary to prepare even a heme group is unknown.

IV. CONFORMATIONAL AND SOLITON SWITCHING

At the heart of molecular electronics is the dual concept of conformational change and charge localization and control. In the following three sections this point will be stressed and we will demonstrate opportunities for using conformational changes (i.e. changes in structural bond lengths and angles) in both the neutral and charged states in MEDs. Per usual nature points the way in providing us with an excellent example of a conformational switch in the photoactivated conversion of cis-retinal to all trans-retinal in the visual system (see Fig. 5). Of particular interest for possible MED application is: 1) The very large conformation change about the 11-12 retinal double bond as indicated in Fig. 5. As suggested earlier [1a], this rotational effect could be used to switch on or off the energy transferred between two excimers (or exciplexes). 2) The charge separation occurring at the double bond upon photoexcitation constitutes an additional important observable phenomenon. This effect has been discussed by L. Salem as the ''sudden polarization'' effect [17] and was postulated in 1970 by W.G. Dauben [18] in order to explain his photocyclization products. 3) As will be discussed later in this section Fig. 5 also provides a mechanism for generating a single charged soliton. A more recent model of the role of proteins in the visual pigments has been given by B. Honig and his associates [19].

One of the most interesting and versatile conformational switching mechanisms appears to be that associated with soliton propagation in a conjugated system (i.e. a system of alternating single and double bonds). Such solitons of special interest here have been discussed in considerable detail previously in this conference by Su [20] and Barrett [21] while Scott [22], Bishop [23], and McLaughlin [24] have treated other varieties of solitons. For our purposes we note simply that the soliton is a nonlinear disturbance which behaves like a pseudoparticle and travels subsonically in a dispersive medium but without being dispersed. The soliton of special concern here is primarily an excited electronic state associated with zero, one, or two electrons occupying a p-π orbital in a conjugated system with the associated soliton charge of +, 0, or - and of spin 0, $\pm 1/2$, 0 respectively. In terms of conformational changes it is important to remember that soliton propagation effectively interchanges the single and double bonds in its path.

Prior illustrations of a soliton switch involved imbedding an organic chromophore in a trans-polyacetylene chain such that they shared bond conjugation [8]. However, such a switch can be

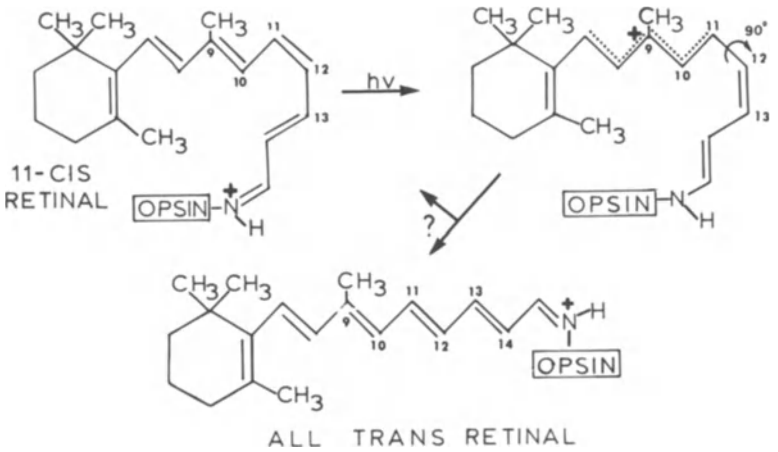


Fig. 5 The photoactivated conformational conversion of 11 cis-retinal to all trans retinal is associated with extensive charge separation strongly assisted by the quaternary imino-nitrogen (see Salem, Ref. 17).

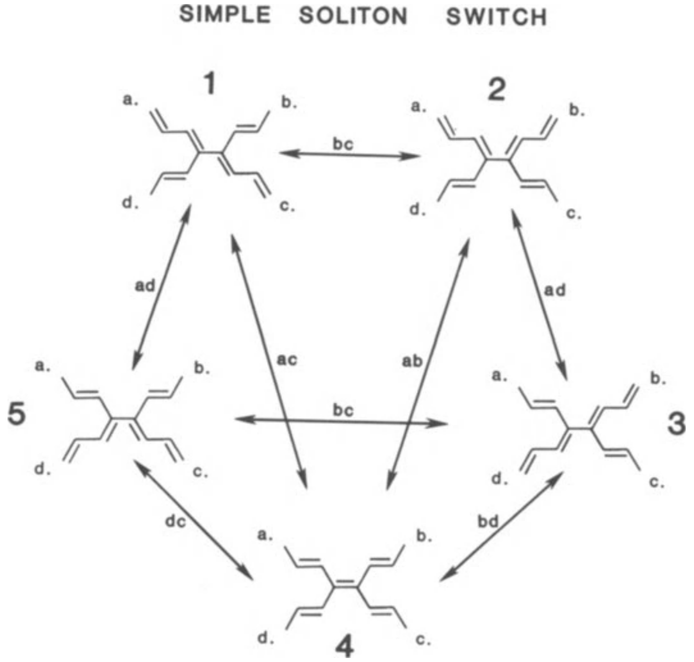


Fig. 6 In this switch the eight arrows indicate the allowed, of fifteen possible, paths. Only the fourth configuration permits paths ac and bd in which the solitons do not change direction.

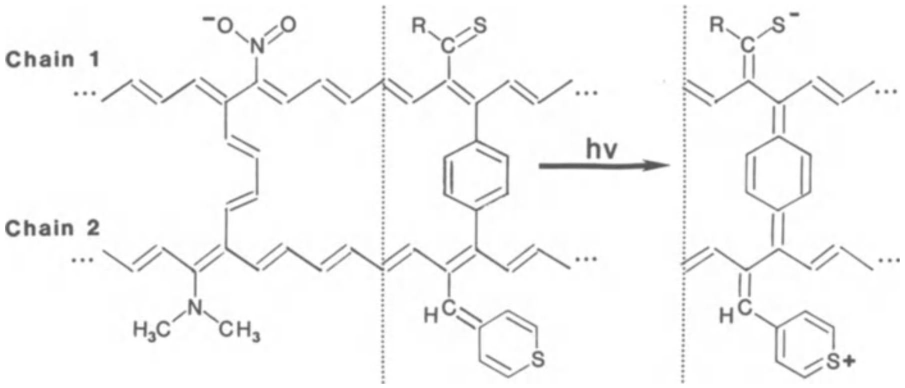


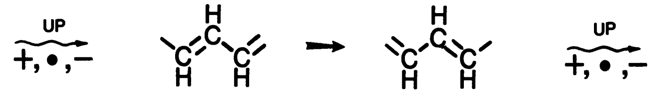
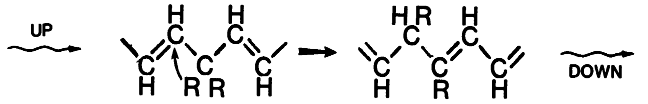
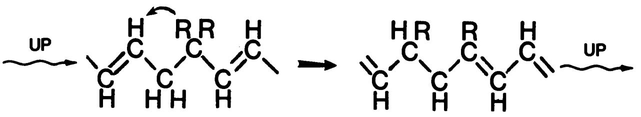
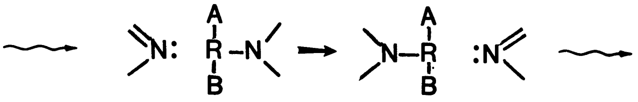
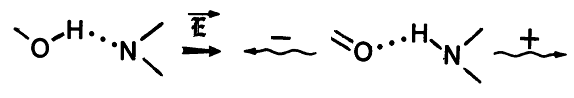
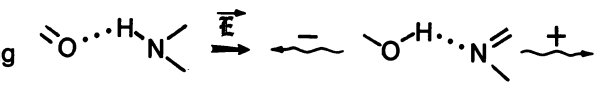
Fig. 7 In this configuration the second sulfur chromophore can be photoactivated but not the first. This soliton gang switch can be readily extended to three trans polyacetylene chains and eight chromophores.

schematically illustrated, Fig. 6, as four communicating lengths of trans-polyacetylene. Of the fifteen conceivable ways of sending a soliton from one end to another end (length) of trans-polyacetylene, seven are prevented by the tetravalency of carbon. The eight allowed soliton paths are indicated in Fig. 6 where a soliton path ab = ba. As a pseudoparticle a soliton's momentum is a constant of motion; this suggests that the primary paths of interest should be represented in Fig. 6 by the direction ac and bd. Only the fourth configuration in Fig. 6 permits both ac and bd soliton propagation paths. However, a soliton path bd changes the configuration to 3 which does not permit an ac soliton path. Hence the bd propagation switches off the ac possibility until a second bd soliton restores the switch to configuration 4. A soliton path, like ac in 3 of Fig. 6, is recognized as being switched off when the conjugation is broken as by two or more successive single bonds.

Fig. 7 illustrates the incorporation of two different chromophores across two trans-polyacetylene chains. While the second chromophore is subjected to photoactivation, as shown, the first chromophore to the right has been switched off by chain 1. A soliton propagating through chain 1 will switch on the first chromophore and switch off the second chromophore; while a soliton in chain two will switch off both chromophores. This concept is readily extended to three chains and eight chromophores with an associated demultiplexing table [8]. By keeping chromophores ~200Å apart to prevent crosstalk and assuming average packing separations of 50 and 100Å in the other directions the packing densities for such gang switches is estimated to be 10^{18} gates/cc.

TABLE 2

NONLINEAR BONDING EFFECTS IN SOLITON PROPAGATION AND GENERATION

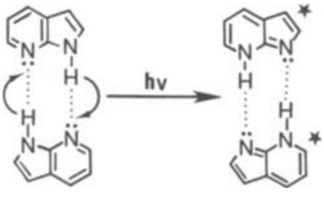
		Ref.
1. Soliton Propagation		20,25
2. Propagation 1,2-shift		25
3. Propagation 1,3-shift		25
4. Mechanical Actuation		26
5. Generation, Proton Tunnelling		8,Sect. VI
6. Generation, Hydride Tunnelling		8,Sect. VI

Soliton propagation in a dispersive medium is necessarily a non-linear effect as shown by Scott, Chu, and McLaughlin [22]. In conjugated systems the non-linearity arises during the conversion of the double bond energy to the single bond in the wave front and the absorption of that energy in the reverse process in the wave tail. Non-linear bonding effects in several postulated soliton phenomena are summarized in Table 2. In each case it will be noted that bond orders are dramatically rearranged (e.g., for the double bond to single bond the bond order goes from 2 to 1, etc.).

The first item in Table 2 shows the bond rearrangement for simple soliton propagation in a short section of trans-polyacetylene. The tetravalency of carbon produces a relationship between direction of soliton motion, the slope of the double bonds and whether its an UP or DOWN soliton [25]. An UP soliton has its center on the upper carbons of the (CH)x chain, a DOWN soliton has its center in the lower carbons. As the soliton moves this UPNESS or DOWNness is maintained because the center of ''kink'' of a moving soliton advances along the chain by two carbons rather than one. Item 2 of Table 2 suggests a soliton can propagate past a carbon sp^3 defect if the 1,2 shift transfer rate is sufficiently enhanced by the approaching soliton [25]. Note propagation of a soliton past this defect changes it from an UP soliton to a DOWN soliton. However, a soliton passing an adjacent pair of sp^3 carbons (item 3) remains an UP soliton. For the soliton mechanical actuator [26] in item 4 of Table 2 the nonlinearity of N-R bonds is indicated by the changes in bond order of 1 to 0. In this actuator device the motion of the R atom by a few tens of an Angstrom drags with it the A and B groups by a similar amount. This mechanical motion is probably sufficient to switch on or off many biological reactions or processes (unspecified). Items 5 and 6 indicate the rearrangement of the hydrogen bond order in the generation of soliton pairs in an electric field by proton and hydride ion transfer respectively [8]. Another example of these processes of soliton generation will be given in Section VI.

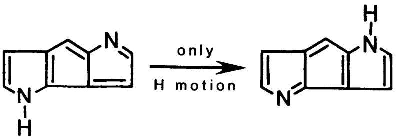
The final effect to be discussed in this section, cooperative particle transfer, while not strictly a soliton phenomena is obviously related to the nonlinearity bonding effect of Table 2 via hydrogen bond changes and to the electric field generation of solitons via proton, etc., tunnelling. In cooperative particle transfer [1b] by the collective action of several functional moieties, a small particle is ''absorbed'' at one site and a second particle is released some distance away, 50Å or greater.

The plausibility of such an effect is suggested by the observation [27] that 7-azaindole exists as a hydrogen bonded dimer in moderately dilute solutions in hexane and when excited at $\lambda = 260$ nm apparently undergoes a double proton transfer as indicated below:



Streak camera studies [28] of the fluorescence spectra of the excited states (*) on the right (a) indicate that the double proton transfer takes place in less than five picoseconds.

Cooperative particle transfer can be illustrated as an extension of the double proton transfer via the use of a special oligomer. In Fig. 8 we wish to transfer a hydrogen atom (proton) from area A at the botton to area B at the top. This is achieved upon photoabsorption by the hydrogen transfer of the H^A hydrogen to the nitrogen in 4-azaindole (I) at the bottom while hydrogens are cooperatively transferred by the intervening tetra-azaindacene derivatives (II) to the moieties above. As the top 4-azaindole moiety receives a hydrogen from the moiety below it can give up its amine hydrogen, H^B, in the 5-member ring to a proton receiver in area B at the top. There are several points to note in this example: (1) The 4 azaindoles on the right side of the equation are both in an excited state (indicated by *) and normally would decay rapidly. (2) The tetraazaindacene moieties (II) on both sides of Fig. 8 are in a ground state in the absence of an electric field, i.e., the proton transfer agent is bimodal with respect to hydrogen coordinate. (3) Note also that the + charge in each II moiety, which is balanced by the anion A- attached to the chain, has shifted by a distance d parallel to the net proton transfer direction. The net charge motion per H is then d* times the number of II moieties. This suggests that the entire process could be controlled by the presence and strength of an electric field as well as by the photon energy. (4) By proper choice of molecules it is possible to both make the end members I and the intervening H transfer molecules II very nearly bimodal and free of major charge transfer; for example consider the three member ring system below in two tautomeric states having a similar valence resonance form.



(5) Finally, we note that this system (Fig. 8) transfers one H atom that must be restored before it can transfer an H atom again, i.e., there is no net material flow.

There are of course several interesting theoretical questions, e.g., (a) in this cooperative particle transfer are all the H atoms transferred simultaneously; or (b) how many members are necessary before the particles are not transferred simultaneously; and (c) in the later case is the particle transfer very similar to a soliton but in an electric or photoelectric field? The answers to these

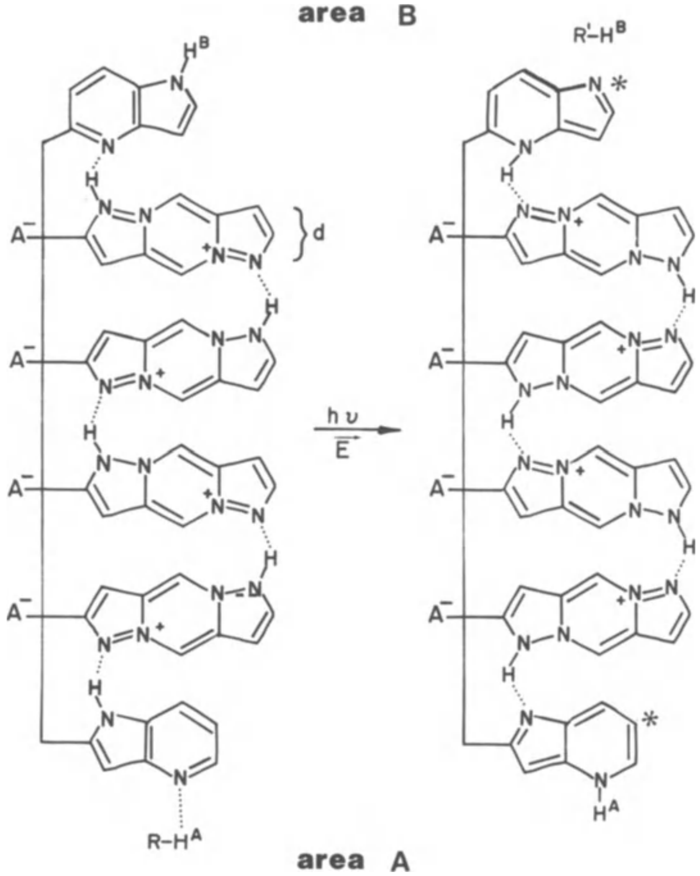


Fig. 8 In this proposed photostimulated cooperative mechanism, a hydrogen is transferred from area A to area B via a series of intermediate transfers which are primarily energy independent. However, note that the end members 4-azaindoles are left in an excited state (*).

questions may of course have relevance to the proton pumps found in biological systems (see Ref. 29).

V. SOLITON VALVES AS ELEMENTARY AUTOMATA

Soliton induced conformational changes in small assemblies of soliton valves can also perform as small discrete

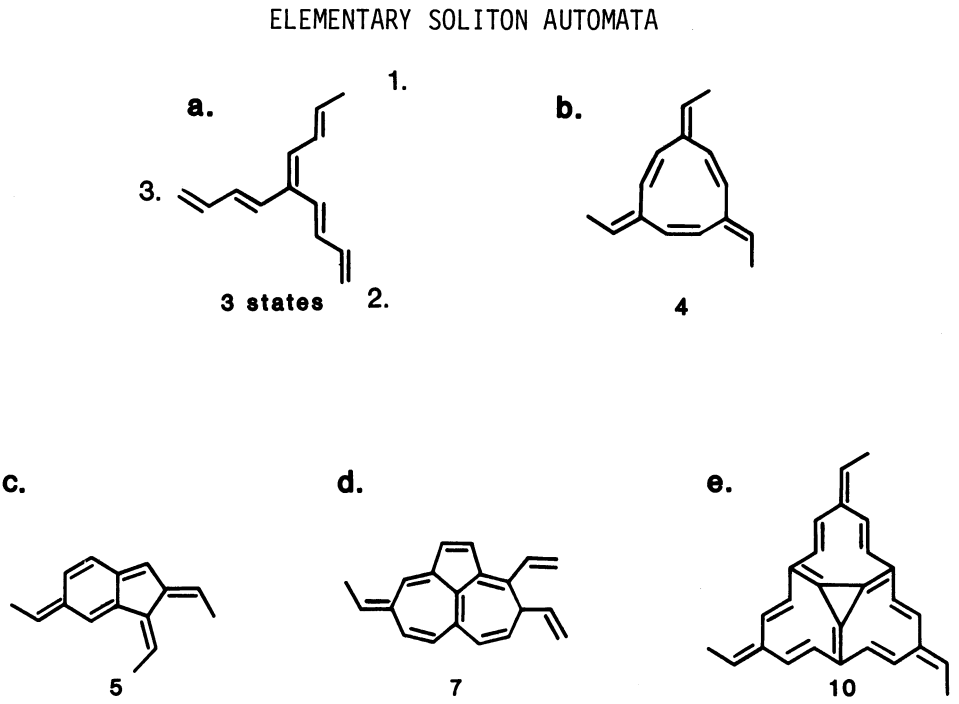


Fig. 9 The valve of Fig. 9a is used in various cyclic configurations, Fig. 9b-e, to obtain different number of states via soliton propagation down their three semi-infinite $(CH)_x$ chains. Strain can be relieved in these configurations if the number of bonds between valves is kept even or odd.

computational systems or automata. A soliton valve may be formed by joining three semi-infinite trans-polyacetylene chains at a single carbon atom [8]. Such a valve is illustrated in Fig. 9a. The position of the double bond on chain 1, is relocated by the passage of the soliton, 12, to chain 2. In this regard the passage of a soliton is like a group operation with the group corresponding to the three-fold rotation group. Note that in Fig. 9a, while 12 and 13 are allowed, 23 is not because of two adjacent single bonds, hence the soliton valve is related but not isomorphous with the ordinary three-way valve.

Configuration of valves as in Fig. 9b-e are of interest in that a different number of states are available in each configuration via soliton propagation. Just as the valve Fig. 9a has three states, a cyclic configuration of three valves with an odd number of bonds between, Fig. 9b, has four states (enumerated in Ref. 4) and forms a soliton propagation group isomorphous with the d_2 point group. The bicyclic configurations of Fig. 9c are listed in Ref. 30 and the 10 different states of Fig. 9e are indicated in Ref. 31 (WMED2).

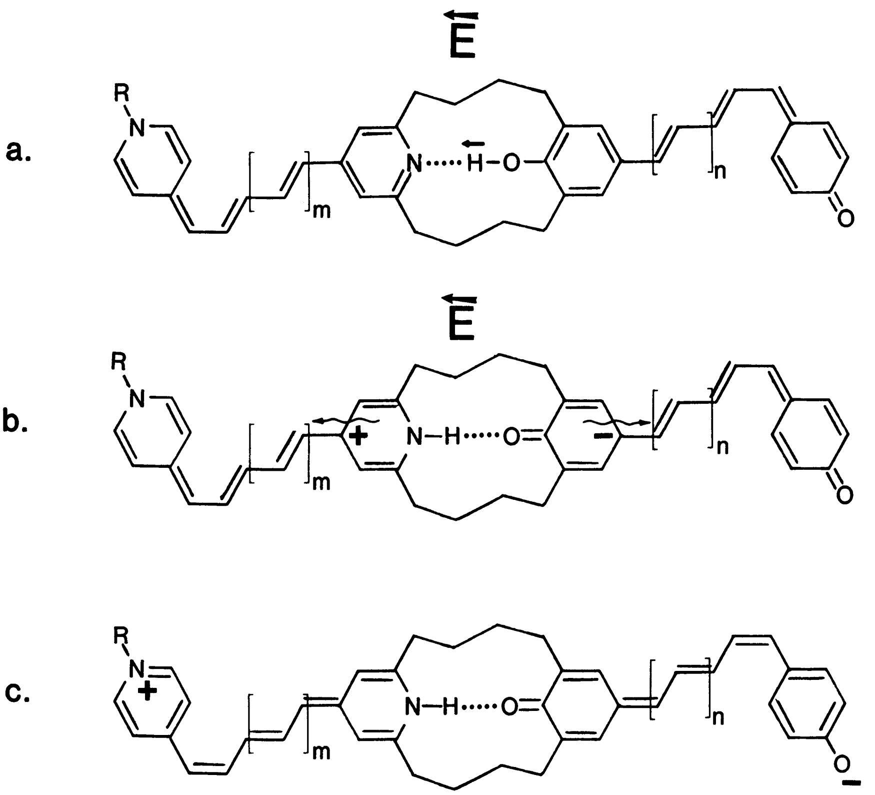
The configuration of valves of Fig. 9 can be considered as very elementary cellular automata (see Ref. 32 for a review article) which are capable of having a number of different states (i.e., 3,4,5,7 and 10) interrelated by soliton propagation. Such valve configuration not only suggests the possibility of computing in other than the binary basis, but also the possibility of using soliton propagation to do sophisticated group operations in a very straightforward fashion. It is noted in addition that M. Groves in following a more traditional approach has illustrated how the elements a and b of Fig. 9 can be used to generate OR and AND gates, and inverters as well as fanout memories (33, WMED2).

VI. SOLITON GENERATION AND VISUAL DATA PARALLEL PROCESSING

It is the two-fold intent of this section to 1) propose a method for testing the generation of soliton pairs by proton tunnelling and 2) then to suggest that this method might be generalized to obtain a color sensitive electro-optical MED camera. This latter idea is then extended to a visual data processor paralleling the data reduction that takes place in the optic nerve.

It was suggested earlier by the author that charged soliton pairs may be generated in an electric field by proton and hydride ion tunnelling [8]. An elaborated version of this method involving proton tunnelling and suitable for testing is offered in Fig. 10. When the electric field E is sufficiently strong the phenol hydrogen will tunnel as a proton to the pyridyl nitrogen (Fig. 10a), generating a charged pair of incipient solitons (Fig. 10b) which will propagate through the polyacetylene chains to the chromophores terminating each chain (Fig. 10c). If a Langmuir-Blodgett film can be made of long chain molecules as indicated in Fig. 11 then proton tunnelling could be readily studied as a function of electric field and incident photon energy by observing either a change in capacitance of the film or a change in the light absorption of the terminating chromophores [31]. Clearly in the presence of a varying electric field such a film could be used as a color sensitive optical detector by noting the capacitance as function of field.

In part soliton generation can be seen as the inverse process of a soliton meeting an sp^3 defect as discussed in Ref. 25 (Table 2, items 2-3). As the hydrogen bonded proton starts tunnelling to the pyridyl nitrogen the effective soliton masses are high and their lengths in the aromatic rings are short



PROTON TUNNELING AND SOLITON GENERATION

Fig. 10 In the presence of an electric field E proton tunnelling would give rise to the generation of a soliton pair, Fig. 10b, which would propagate to the terminal chromophores. For the trans polyacetylene chains we assume m ~ n > 6.

(Fig. 10b), however as they enter the trans-(CH)x chains they become more typical solitons until they arrive at the chromophores.

The testing scheme of Fig. 11 can be made position sensitive by either reticulating the conducting capacitor films into small individual elements or by not terminating all the polyacetylene chains with chromophores but letting some of them penetrate the conducting films to computing units or automata as mentioned in the previous section. We note that if one is able to synthesize an MED as by the modular method of section II then one should be capable of simultaneously synthesizing two and then

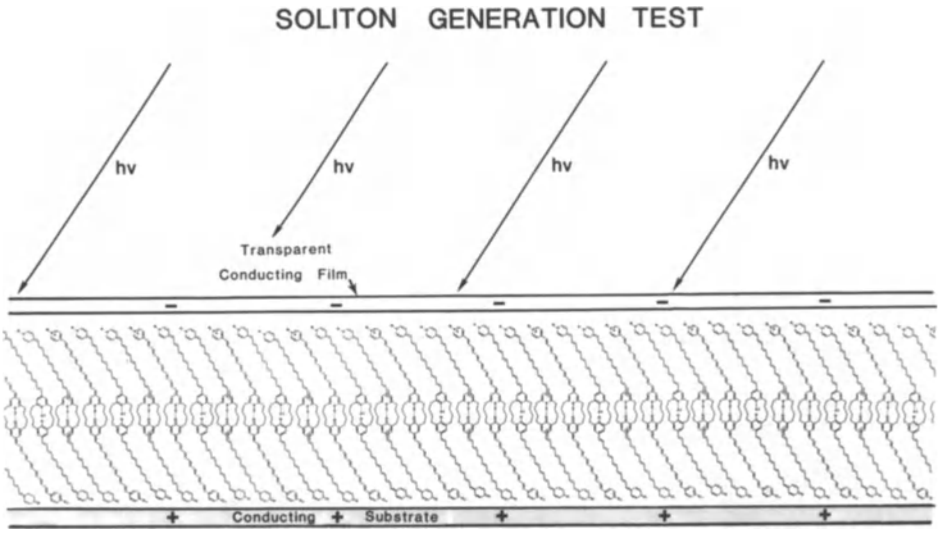


Fig. 11 Proton generation of charged solitons can be tested by observing capacitance changes as a function of field and photon energy in the sandwich film arrangement indicated here.

three-dimensional arrays of processing automata with localized memory capacity.

Such a 3-d array of automata attached to color sensitive optical detector is then indicated in Fig. 12 and in a naive way has a strong parallel to the retina and the optic nerve. We note that moving in the dimension away from the optical detector the data analyses becomes increasingly sophisticated. Using parallel processing data analysis could be extremely fast even using solitons as the data transport signal. The amazing speed with which athletes interpret visual data confirms this even though neural signal speeds are subsonic.

VII. LITHOGRAPHY AT THE MOLECULAR LEVEL

The increasing need of the semiconducting industry for improved lithography at the nanometer scale, as well as even more severe requirements for the production of future MED devices, suggests that a new approach to lithography is required. Such a requirement provides another synergistic biotechnical opportunity. This opportunity would arise from the development of a technology in which structures are built up from the molecular level in contrast to the current lithographic process of scaling down. For example, it has been suggested

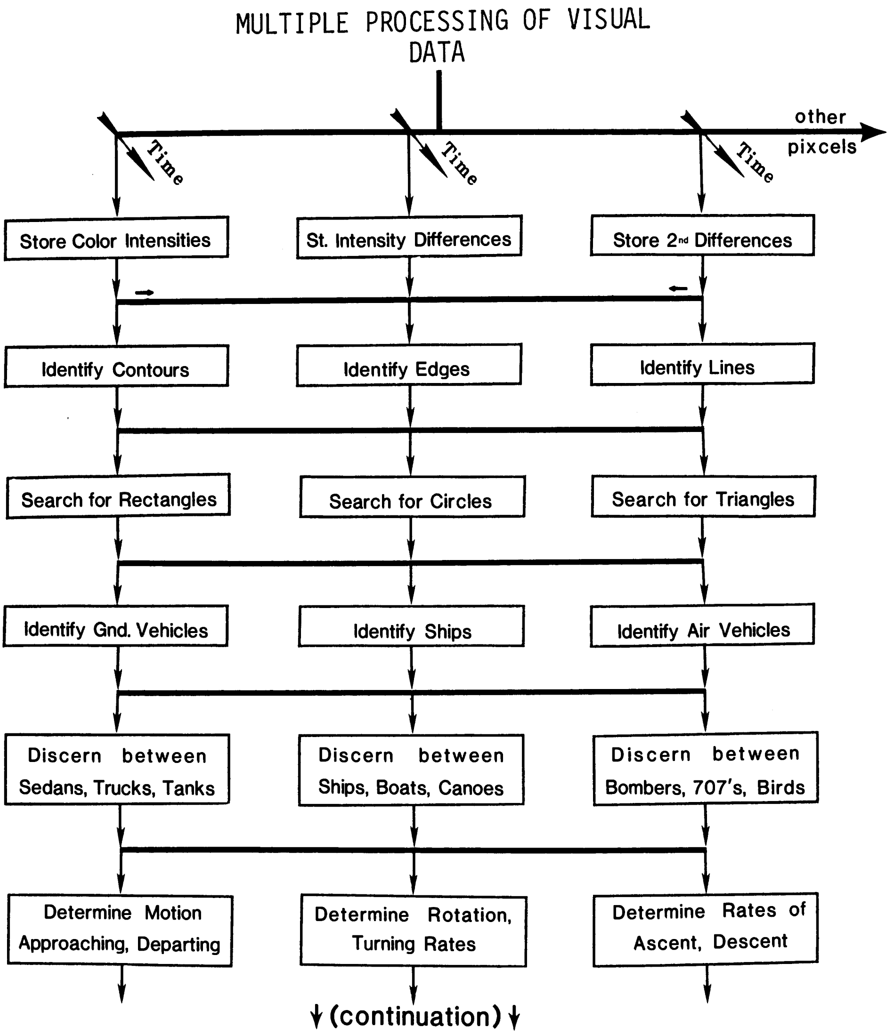

MULTIPLE PROCESSING OF VISUAL DATA
other pixcels
Time
Time
Time
Store Color Intensities
St. Intensity Differences
Store 2nd Differences
Identify Contours
Identify Edges
Identify Lines
Search for Rectangles
Search for Circles
Search for Triangles
Identify Gnd. Vehicles
Identify Ships
Identify Air Vehicles
Discern between Sedans, Trucks, Tanks
Discern between Ships, Boats, Canoes
Discern between Bombers, 707's, Birds
Determine Motion Approaching, Departing
Determine Rotation, Turning Rates
Determine Rates of Ascent, Descent
(continuation)

Fig. 12 A three dimensional array of parallel processors could be linked to individual pixcels to provide a rapid interpretation of visual data.

earlier [26] that superstructures might be formed on silicon or GaAs by using a large flat molecule B, that can be deposited expitaxially as in Fig. 13a. If these B molecules are designed so that their intermolecular forces enable them to form a non-area filling monolayers as in Fig. 13b, then a regular pattern of voids, A, can be formed. These areas, A, can be filled in via a number of techniques such as electrochemical transport, cytochemical staining, sputtering, or vapor deposition.

MOLECULAR EPITAXIAL SUPERSTRUCTURING

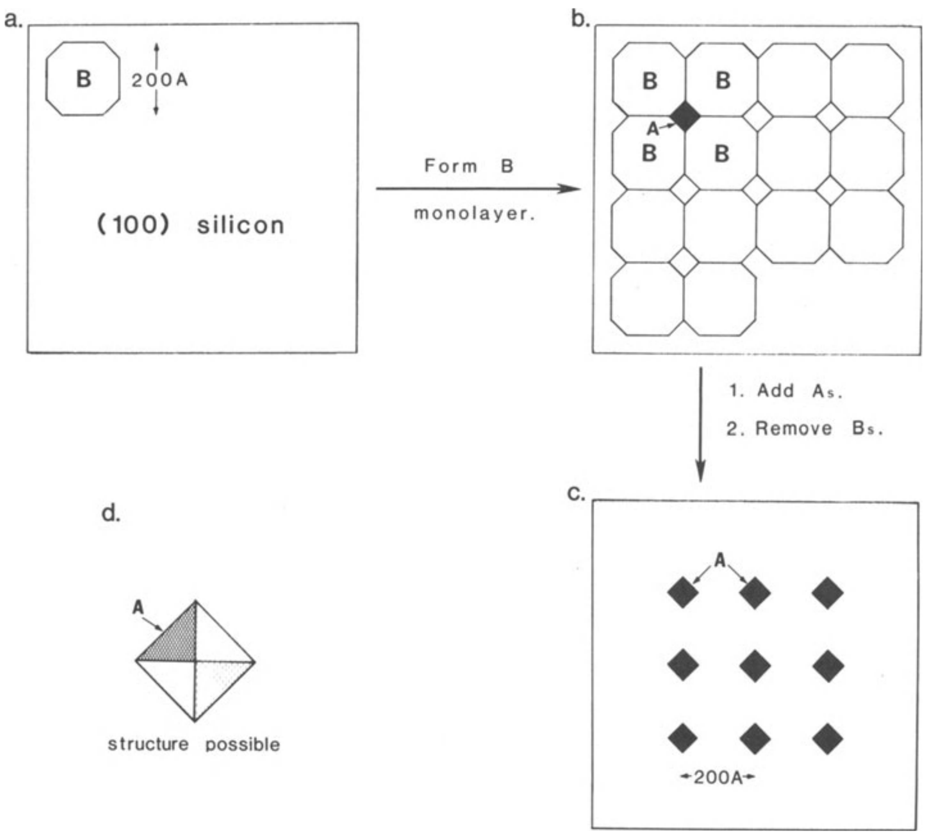


Fig. 13 Epitaxial deposition of large B molecules on single crystals can result in lithography at the molecular level if the B molecules are designed to form the correct pattern. The pattern indicated in Fig. 3c is obtained by filling in the pattern holes of Fig. 3b and then removing the B epitaxial layer.

Thereafter the B molecules can be removed, for example, by photolytically decomposing them as suggested in Fig. 14. This, then, leaves the superstructure of filled A sites as is indicated in Fig. 13c.

If protein B molecules were suitable then presumably they could be bioengineered using recombinant DNA techniques as

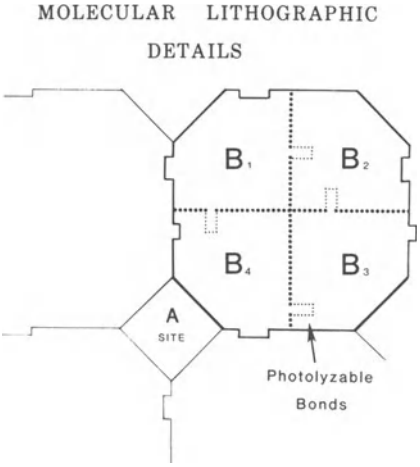


Fig. 14 Using recombinant DNA techniques the large B molecule could be designed to interlock at the edges and be readily broken down for easy removal after A group deposition.

recently described by Ulmer [5,15] to have strong B-B edge interactions and photolysable bonds. As if to substantiate the above possibility more recently Ulmer [16] has discussed self-organizing protein monolayers. Ulmer paid particular attention to the influenza virus neuraminidase which forms a two dimensional monolayer (see Fig. 15) very much as suggested in Fig. 13. Structural data indicates that the box-shaped head

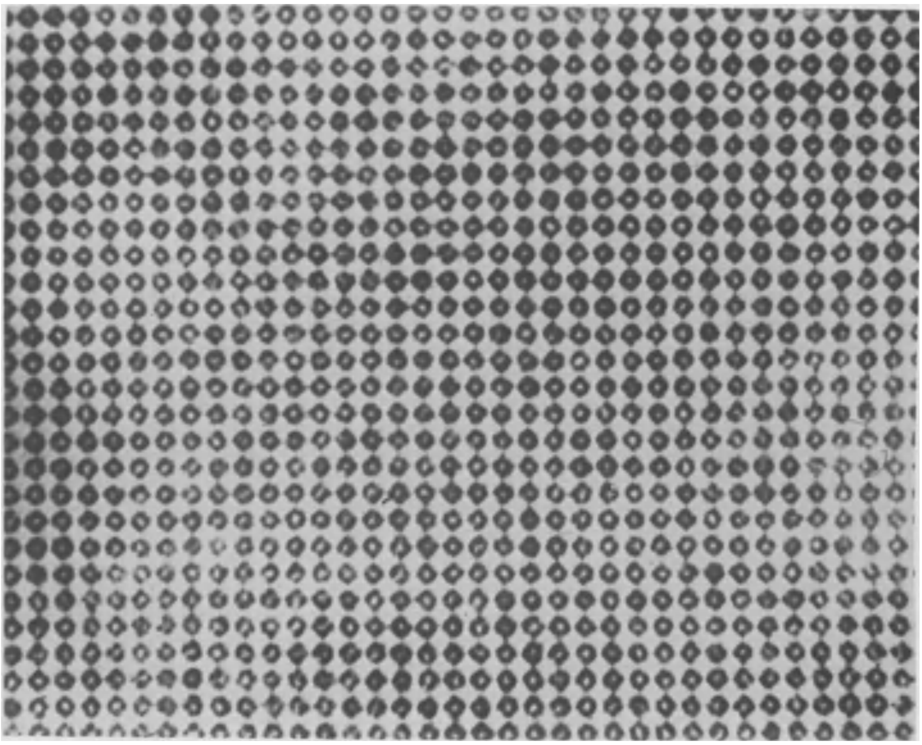


Fig. 15 This two dimensional square array of neuraminidase heads on 100 A centers is strikingly similar to Fig. 13b. The electron micrograph was kindly supplied by D.C. Wiley, Biochemistry, Harvard and N. Wrigley and E. Brown, Nat. Inst. of Med. Research, Millhill, London.

100 x 100 x 60A, is a tetramer not unlike the B molecules of Fig. 14 [34].

Contrarywise the development of tight expitaxial coatings for semiconductor products could have applications in prosthetic implants as substrates for compatible biological coatings.

VIII. ORGANIZING MOLECULES IN L-B FILMS

In this section we wish to emphasize how the concepts from two different fields can combine to contribute to a third technology. In particular we combine a biological approach to understanding ion conductance through lipoid bilayers [35] with materials developed to study energy and electron transport

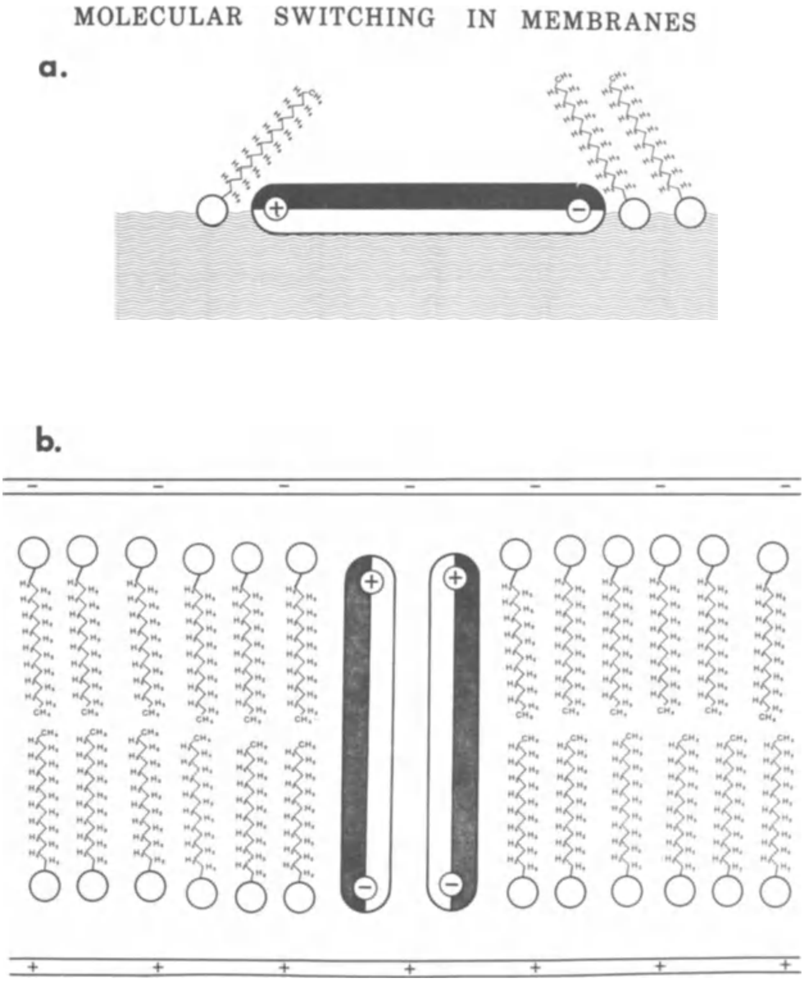


Fig. 16 To explain ionic conduction through lipoid bilayers Baumann et al. [35] suggested that rod shaped protein molecules with one side hydrophilic and the other hydrophobic as in Fig. 16a, are drawn into bilayers in the presence of an electric field to form a hydrophilic channel (Fig. 16b).

through Langmuir-Blodgett (L-B) films [36]. What we obtain is a technique which might be useful for developing two-dimension structure in L-B and other films. Such films have potential application in MED, semiconductor and medical technology.

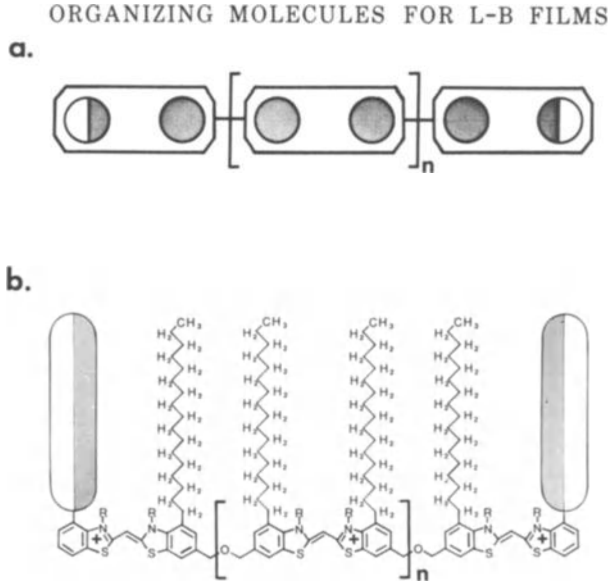


Fig. 17 Combining the ideas of Kuhn and Baumann, the author proposes that a series of molecules can be prepared that uses weak van der Waals and dipole forces to organize L-B films. The stylized top view of Fig. 17a is used in the next two figures.

In order to explain voltage dependent ion current switching in lipoid bilayers, Baumann, Easton, Quint, and Johnson, have suggested that biphilic protein rods cluster in membranes to form hydrophilic channels [35]. Such a protein rod is indicated in Fig. 16a on the aqueous surface. However, in the presence of an electric field such rods enter the membranes to form channels by the clustering of 3-5 rods.

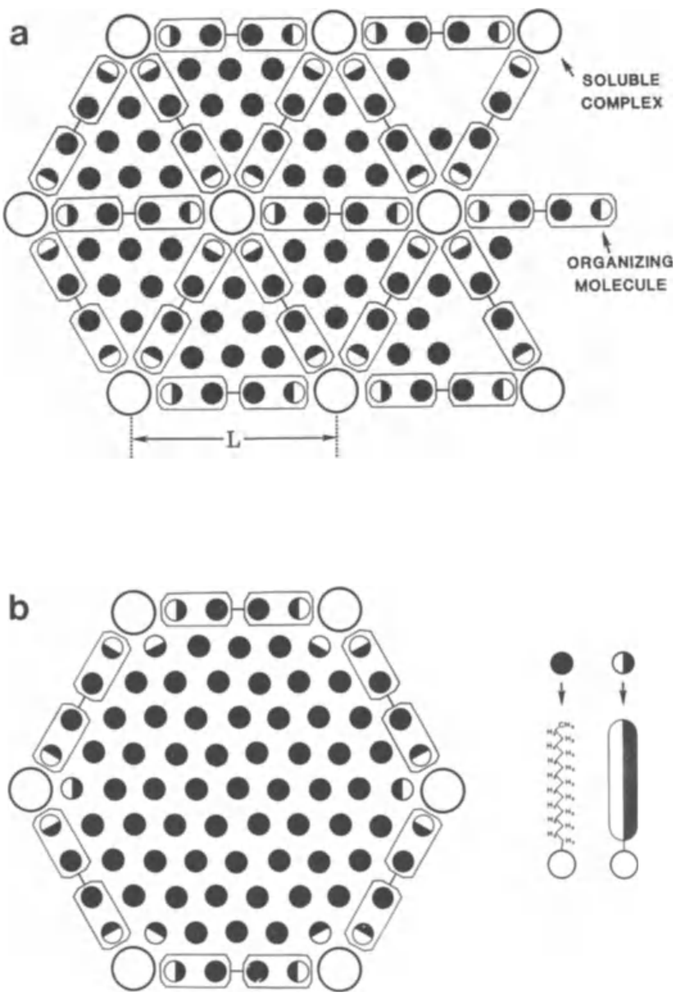

ORGANIZING L-B FILMS

Fig. 18 The use of ''organizing'' molecules and composition variations should lead to 2-d periodic structures of different cell sizes. If the unit cell length in Fig. 18a is L then for Fig. 18b it is 3 L.

In order to stabilize the presence of electron acceptors and cyanine dye molecules in L-B films H. Kuhn [36] has attached two long chain aliphatic groups to each functional molecule. By combining these ideas, molecules suitable for organizing L-B films are possible as suggested in Fig. 17. The side view in Fig. 17b shows n+2 cyanine dye molecules linked by ether bridges. Of particular interest note that the terminal members

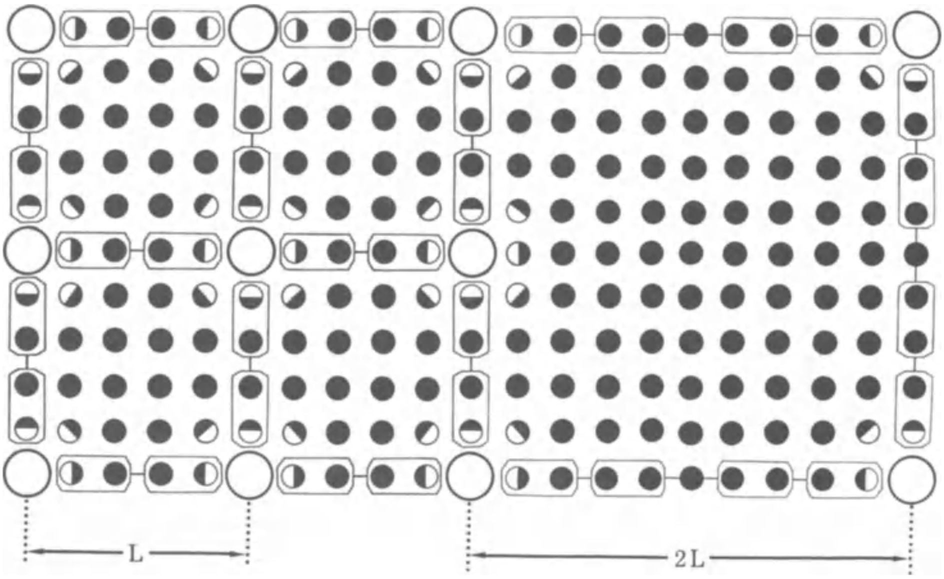


Fig. 19 By changing the organizing molecule in addition to compositional variations the unit cell size might be doubled as indicated here for a square lattice.

contain the biphilic rods discussed by Baumann et al. [35]. These rods then permit the organization of L-B films by clustering as in Fig. 18. Three points are of interest: 1) Fig. 18 indicates the presence of a hydrophilic soluble complex. Such a complex might be composed of several small molecules that might later be dissolved out to form a skeletonized L-B film with periodic pattern of holes. 2) Such a film placed on a substrate may provide a regular array of ion conducting paths for a variety of lithographic and other purposes. 3) As indicated by comparing Figs. 18a and 18b it is clear that the effective cell size can be varied by the film composition. The unit cell edge in Fig. 18b is 3 times that in Fig. 18a.

Fig. 19 illustrates another technique for obtaining a new and larger cell size. This is possible by employing longer organizing molecules (that is, n > 0). Such a periodicity change provides the chemists with an equivalent to fractal relationships (pattern repeats at a different scale) obtained in cellular automata [30, 32].

Finally, we note that stabilized (i.e. polymerized) pinhole-free L-B films may play, soon, an important role in the fabrication of Josephson junctions and metal-insulator semiconductor field-effect transistors (MISFET). It seems certain that stabilized films with periodic features as in Figs. 18 and 19 will play important roles in biological and medical devices as well.

IX. LOOKING AHEAD

In this final section we would like to take note of the last four items (6-9) of Table 1 listing possible synergistic MED-biotechnical opportunities.

A hybrid MED-semiconductor detector (item 6, Table 1) based on the field effect transistor (FET) may be the first MED device to be realized. In the FET electrons tunnelling from the source to the drain in a small volume of silicon near to the silicon-oxide metal gate is controlled by the gate potential (for a review of CHEMFETs, see Ref. 37). By replacing the oxide layer and metal gate with thin, perhaps monomolecular, layer of ''molecular'' MED active elements, the potential in the silicon can be controlled and hence the electron tunnelling as well. For example, by using a molecular layer of chromophores, a MED photodetector has been proposed that makes use of the changing dipole field [38]. In such devices (hybrid FETs) it would be very important to 1) stabilize the MED-Si interface against oxidative attack, and 2) to place the MED component as near to the Si surface as possible. Since the GaAs surface is less oxygen sensitive and has high electron mobility, GaAs may be the semiconducting component of the future. Currently, variations of FET detectors are being investigated in which monoclonal antigen and enzyme fragments are bonded to the gate surface area [37]. While initially such devices are sensitive to biological reactions their long term stability is currently rather poor.

Another important opportunity is that provided by combining MED and biotechnology in prosthetic devices (item 7, Table 1). These might range from control microprocessors for artificial limbs to ultimately, artifical sensors with visual and aural functions. As a particular example, consider the uses for a ''chemical factory on a chip'' that might be small enough to be implanted in the blood stream [39]. For diabetics, as an example, one side of the chip might incorporate sensors for blood sugar, ph, etc., whatever is needed for analysis of current state of being. The other side might contain a biochemical factory for the production of human insulin, which factory could be controlled by a microprocessor upon analysis of the input data. While each such device need be individually tailored for immunological reasons, the humanitarian and social impact of such a technology is enormous.

From the work of J.J. Hopfield [40] it appears likely that content addressable memory (CAM) systems (item 8, Table 1) may be obtained from the study of neural systems. Practical CAM systems have important implications, not only in terms of the indexing of archival memories, but also in terms of error coding, a topic of certain interest in MED technology. Let us speak of memory bits of information per memory units (MU), where memory units can be readily identified as discrete gates, etc., in digital systems but it is not clear what constitutes a memory unit in the human nervous system. For example, data might be stored on a simple tape system in terms of 8 bits of information plus one parity bit in 9 MU on the tape. A single error in the eight data bits can be detected by summing them and comparing their evenness or oddness with the parity bit. For larger data words, that is 28 bits in 32 MU, Hamming coding can be used to correct a single error and detect double errors. At the far end of this scale is associative or CAM systems like the human nervous system where 100 bits might be stored in thousands of MUs. Using a simple synaptic model Hopfield [40] has developed a memory system (CAM) where very partial information permits full data recovery by a reiterative process. Pictorially the partial data locates one on the sides of a valley and the interaction leads one to the low point of the valley and full data recovery. It seems likely that all of the models, parity bit, Hamming coding, and CAM, are part of a to be generalized mathematical formula as a function of information bit/memory unit. It is also clear that MED devices will be somewhere on the scale between Hamming coding and the CAM model of Hopfield.

In a similar vein and as a final example (item 9, Table 1) of a biotechnical opportunity, we want to emphasize the point recently made by P.E. Rapp [41] that the study of the human nervous system is appropriate both for those interacted in hierarchical control of parallel processors and those interested in the reliability of high density devices; these are in addition to its humanitarian values. To stress the first point we note that modern aircraft currently contain as many as 50 interlinked control microprocessors and that future aircraft will contain as many as 250 [42]. Similarly, for the safety of all those who fly, overall system reliability must be retained even as individual components becomes smaller, more dense, and more certain to be individually unreliable.

X. CONCLUSION

The few examples pointed to here are but a small sampling of those available. Even so it is evident that the interdisciplinary and coherent development of molecular electronics and biotechnology provide rewards and challenges, scientific, technologic, and humanitarian, sufficient to be of national concern.

REFERENCES

1. a. F.L. Carter, ''Problems and Prospects of Future Electroactive Polymers and 'Molecular' Electronic Devices,'' in the NRL Program on Electroactive Polymers, First Annual Report, Ed. L.D. Lockhart, Jr., NRL Memorandum Report 3960, p. 121 (1979): ''Further Considerations on 'Molecular' Electronic Devices,'' Second Annual Report, Ed. R.B. Fox, NRL Memorandum Report 4335, p. 35 (1980).
2. A. Aviram and M.A. Ratner, Chem. Phys. Letters, 29, 277 (1974).
3. a. F.L. Carter, Ed., ''Proceedings of the Molecular Electronic Device Workshop,'' held 23-24 March 1981, NRL, Washington, DC, NRL Memorandum Report 4662, October 1981; also as ''Molecular Electronic Devices,'' published by Marcel Dekker, N.Y., N.Y., Dec. 1982; b. 2nd International Workshop on Molecular Electronic Devices held 13-15 April 1983, NRL, Washington, DC, Marcel Dekker, Inc., in press.
4. F.L. Carter, ''The Chemistry in Future Computers,'' in Computer Applications in Chemistry. Proceedings of the 6th International Conference on Computers in Chemical Research and Education,'' Eds. S.R. Heller and R. Potenzone, Jr., Elsevier Scientific Publishing Co., Amsterdam (1983).
5. K.M. Ulmer, ''Biological Assembly of Molecular Ultracircuits,'' in Ref. 3a.
6. F.L. Carter, ''Molecular Level Fabrication Techniques and Molecular Electronic Devices,'' plenary address at the 1983 International Symposium on Electron, Ion, and Photon Beams, held 31 May-3 June 1983, Los Angeles, CA, J. Vac. Sci. and Tech., in press.
7. J.W. Milliken, ''Metallic Covalent Polymers: Derivatives of (SN)x and (CH)x,'' Chemistry, Ph.D. Dissertation, Univ. of Pa., 1980.
8. F.L. Carter, ''Conformational Switching at the Molecular Level,'' in Ref. 3a, p. 51.
9. F.L. Carter, ''Electron Tunnelling in Short Periodic Arrays,'' in Ref. 2a, p. 344; and NRL Memorandum 4717.
10. E.A. Pschenichnov, Soviet Physics - Solid State, 4, 819 (1962).
11. J.R. Barker, ''Quantum Ballistic Transport and Tunnelling in Molecular Scale Devices,'' in Ref. 3b.
12. S.W. Kirchoefer, R. Magno, K.L. Davis, R.L. Schmidt, and J. Comas, ''Resonant Tunnelling in Molecular Beam Epitaxially-Grown Semiconductor Structures,'' in Ref. 3b.
13. F.L. Carter, ''Studies of Tunnelling in Short Periodic Arrays,'' in Ref. 3b.

14. P.M. Kuznesof, K.J. Wynne, R.S. Nohr and M.E. Kenney, J.C.S. Chem. Comm. 121 (1980); R.S. Nohr, P.M. Kuznesof, K.J. Wynne, and M.R. Kenney, J. Am. Chem. Soc., 103, 4371 (1981).
15. K.M. Ulmer, Science, 219, 666 (1983).
16. K.M. Ulmer, ''Self-Organizing Protein Monolayers as Substrates for Molecular Device Fabrication,'' in Ref. 3b.
17. L. Salem in ''Excited States in Organic Chemistry and Biochemistry,'' Eds. B. Pullman and N. Goldblum, D. Reidel Publishing Co., Dordrecht, Holland, 1977, p. 163.
18. W.G. Dauben and J.S. Ritscher, J. Am. Chem. Soc. 92, 2925 (1970).
19. B. Honig, ''Theoretical Aspects of Photoisomerization in Visual Pigments and Bacteriorhodopsin in Biological Events Probed by Ultrafast Laser Spectroscopy,'' Ed. R.R. Alfano, Academic NY, NY, 1982, p. 281; ''Visual Pigments: A New Type of Molecular Switching Device,'' in Ref. 2a,p. 137; B. Honig and T.G. Ebrey, Methods Enzymology, 88, 462 (1982).
20. W.P. Su, ''Nonlinear Excitations in Polyacetylene,'' this conference; W.P. Su, J.R. Schrieffer and A.J. Heeger, Phys. Rev. Lett., 42, 1698, (1979): Phys. Rev. B22 2099 (1980).
21. T.W. Barrett, ''Soliton Propagation in One-Dimensional Organic Conductors,'' this conference.
22. A.C. Scott, ''Solitons and Bioenergetics,'' this conference; also A.C. Scott, F.Y.F. Chu, and D.W. McLaughlin, Proc. IEEE, 61, 1443 (1973).
23. A.R. Bishop, ''Solitons in Biological and Synthetic Polymers: Some Comparisons and Contrasts,'' this conference; also D. Campbell and A.R. Bishop, Phys. Rev. B24, 4859, (1981); Nuclear Phys., B200 (FS4), 297 (1982).
24. D.W. McLaughlin, ''Solitary Waves as Fixed Points of Infinite Dimensional Maps,'' this conference.
25. F.L. Carter, ''Solitons and SP^3 Defects in Trans-Polyacetylene,'' in Polymers Electroactifs, Eds. P. Bernier, and B. Payet, CNRS, at Ecole d'Hiver, Font-Romeu, Jan. 1982, Vol. 1, p. 146.
26. F.L. Carter, ''Toward Computing at the Molecular Level,'' in Microelectronics-Structure and Complexity, Ed. Raymond Dingle, Plenum Press, NY, 1983, in press.
27. C.A. Taylor, M.A. El-Bayoumi, and M. Kasha, Proc. Nat'l. Acad. Sci. U.S., 63, 253 (1969).

28. K.B. Eisenthal, K. Gnadig, W. Hetherington, M. Crawford, and R. Micheels, ''Picosecond Laser Studies of Electron-Hole Interactions and Double Proton Transfer,'' in Picosecond Phenomena, Proc. of First International Conference on Picosecond Phenomena, Eds., C.V. Shank, E.P. Ippen, and S.L. Shapiro, Springer-Verlag, Berlin, Heidelberg, New York, 1978, p. 34.
29. R.R. Birge, ''Molecular Dynamics of the Proton Pumping Photocycle in Light-Adapted Bacteriorhodopsin,'' this conference.
30. F.L. Carter, ''The Molecular Device Computer: The Point of Departure for Large Scale Cellular Automata,'' presented at the Workshop on Cellular Automata, CNLS, Los Alamos National Laboratory, Los Alamos, N.M., 7-11 March 1983, Physica D, in press.
31. F.L. Carter, ''Soliton Switching and Its Implications for Molecular Electronics,'' in Ref. 3b.
32. S. Wolfram, ''Statistical Mechanics of Cellular Automata,'' Rev. Mod. Phys., July 1983, in press.
33. M.P. Groves, ''Dynamic Circuit Diagrams for Some Soliton Switching Devices,'' in Ref. 3b.
34. J.N. Varghese, W.G. Laver, and P.M. Colman, Nature, 303, 35 (1983).
35. G. Baumann, G. Easton, S. Quint, and R. Johnson, ''Molecular Switching in Neuronal Membranes,'' in Ref. 3b.
36. H. Kuhn, a. Thin Solid Films, 99, 1 (1983), b. ''Self-Organizing Molecular Electronic Devices,'' in Ref. 3b.
37. J. Jonata and R.J. Huber, ''Chemically Sensitive Field Effect Transistors,'' in Ion-Selective Electrodes in Analytical Chemistry, Vol. 2, Ed. H. Freiser, Plenum Press, NY (1980), p. 107.
38. F.L. Carter, ''Photoactivated Field Effect Transistor,'' Invention Disclosure, Navy Case No. 66,931, 1982.
39. J.H. McAlear, Gentronix-Lak Inc., private communication, 1981.
40. J.J. Hopfield, Proc. National Acad. Sci., USA, 79, 2554 (1982).
41. P.E. Rapp, ''Reliability in High Density Devices: Possible Lessons from Neural Systems,'' in Ref. 3b.
42. B.A. Zempolich, ''Molecular Electronic Devices - A System User's Perspective of Opportunities and Challenges,'' in Ref. 3b.

NONLINEAR ELECTRODYNAMICS IN CELLULAR ELECTROCHEMISTRY

RADIOFREQUENCY RADIATION AND CHEMICAL REACTION DYNAMICS

Richard A. Albanese and Earl L. Bell

USAF School of Aerospace Medicine

Brooks AFB, San Antonio, Texas, 78235, USA

The goal of our research is to understand the effects of modulated radiofrequency radiation on chemical reaction dynamics. Our work is motivated by reports of effects of modulated radiofrequency radiation on electroencephalogram activity (Takashima et al., 1979; Bawin et al., 1973; Servantie et al., 1975) and by reports addressing tissue calcium handling (Adey, 1979; Luben et al., 1982). We have specifically attempted to find the simplest chemical system that might demonstrate field modulation effects and non-monotone responses to increasing field amplitudes.

We have studied an enzyme catalyzed irreversible chemical reaction occurring in a volume V. The reaction is designated X→C and occurs with rate k(T) given by the enzyme temperature - activity curve.

$$k(T) = \frac{\kappa A_0 \exp(-\psi/t)}{1 + (k_1/k_2)\exp[-(\psi_1-\psi_2)/T]} \qquad \text{Eq.1}$$

In the above equation the following definitions apply:

κ = the stoichiometric coefficient for the reaction X→C as catalyzed by the active form of the enzyme.

A_0 = the total concentration of enzyme available, including active and denatured forms.

κ_1 = the stoichiometric coefficient for the reaction, active enzyme → denatured (inactive) enzyme.

κ_2 = the stoichiometric coefficient for the reaction, inactive enzyme → active enzyme.

ψ = activation energy in degrees Kelvin for the reaction X→C

Ψ_1 = activation energy corresponding to κ_1.

Ψ_2 = activation energy corresponding to κ_2.

T = temperature in degrees Kelvin.

This equation is derived in the appendix.

For analytical and numerical work we have found it best to write equation 1 as:

$$k\ (T) = \frac{k_{max}\ [\beta/(\beta-\psi)]\exp[-\psi(1/T - 1/T_{max})]}{1\ +\ [\psi/(\beta-\psi)]\exp[-\beta(1/T - 1/T_{max})]} \qquad \text{Eq.2}$$

In this equation, T_{max} is the temperature at which k(T) assumes its largest value, k_{max}, and $\beta = \Psi_1-\Psi_2$.

We assume that the reactant X is freely diffusible across the boundary of the volume V and thus the flux across the boundary is given by the expression:

$$\text{flux} = h(X_e-X)$$

where X_e is the exterior concentration of the reactant X. For the reactant X we may then write the following equation:

$$\dot{X} = h(X_e-X) - k(T)X \qquad \text{Eq.3}$$

We also treat the balance of thermal energy in the volume V. We assume that heat is freely diffusible across the boundary of V with the flux given by

$$\text{heat flux} = \ell(T_e-T)$$

where T_e is the exterior temperature. The heat of the reaction X→C is ΔH so that $\Delta Hk(T)X$ is the heat production (or consumption) due to the chemical reaction. We assume that the system absorbs microwaves and as a starting point have studied solvent heating equal to $1/2\ \sigma E^2\sin^2(2\pi ft)$ where σ is the material conductivity, E is the carrier wave amplitude, and f is modulation frequency (Hz). Combining terms we obtain:

$$\dot{T} = \ell(T_e-T) + \Delta Hk(T)X + 1/2\ \sigma E^2\ \sin^2\ (2\pi ft)$$

The exterior temperature T_e is assumed unperturbed by the radiation field.

Equations 2,3, and 4 are descriptive of an experiment one could perform in vitro. They may also approximate conditions in a tissue if that tissue is perfused by blood or bathed in fluid at constant temperature.

Examining equations 2,3, and 4 the mathematically oriented reader may well immediately suspect that the chemical system under study is capable of exhibiting frequency modulation effects and amplitude windows. A modulation frequency resonance is feasible since this is a second order differential system, albeit non-linear. The non-monotone nature of the dose-response relationship could result from the maximum in the enzyme temperature-activity curve.

We calculated values of reactant concentration (X), time rate of change of reactant concentration $(\dot{X})$, temperature (T), time rate of change of temperature $(\dot{T})$, and time rate of change of product formation $(\dot{C}=k(T)X)$ for given time points and various values of modulation frequency f and field strength E. Equations 3 and 4, as a coupled system of differential equations with appropriate initial conditions, served as a basis for the calculations. Carrier wave frequency was fixed at 400 MHz. The system constants used in the analysis are set out in Table I, and Figure 1 displays the temperature-activity curve determined by the values of k_{max}, ψ, β and T_{max} given in Table I. The constants used were chosen both to evoke the desired system behavior and to be at least potentially realizable in practice.

The differential equation solver, DGEAR in the IMSL system (IMSL Library Reference Manual. Edition 8, Volume 1.) was our primary analytical tool. In looking at our numerical results we focussed on the average rate of product formation as it differed from the rate without radiation. The arithmetic average differential, symbolized $\Delta\dot{C}_{ave}$, based on 200 computed values of $\dot{C}=k(t)X$ equispaced in time, for each modulation frequency and field strength pair, seemed most biomedically meaningful to us as this quantity represents the mean shift in metabolite reduction that would be expected due to the solvent heating effect of radiation. The arithmetic average computed is an accurate numerical approximation to the time average of $\Delta\dot{C}_{ave}$.

Our numerical work is summarized in Figures 2,3, and 4 below. Exquisite sensitivity of product formation rate to modulation frequency is noted. In the curves showing modulation effects at 50 V/meter and 100 V/meter, the solvent temperature perturbations due to radiation were very slight, being less than 0.01°K. Figure

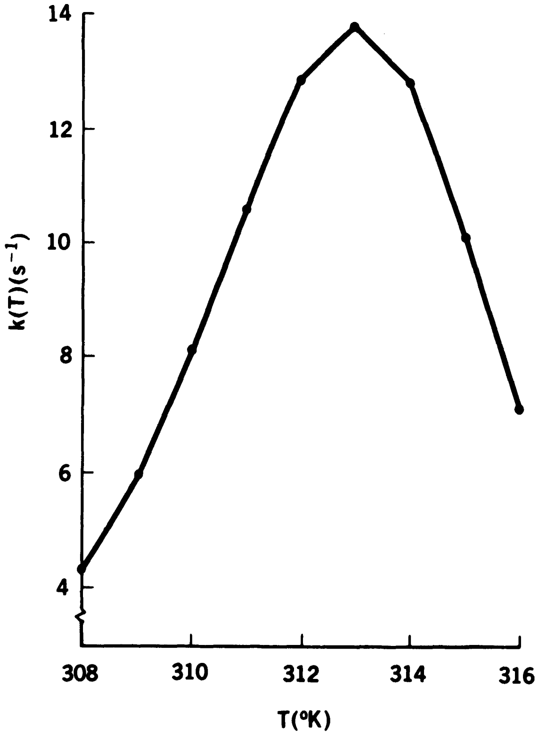


Fig. 1. Influence of temperature on the rate of reaction.

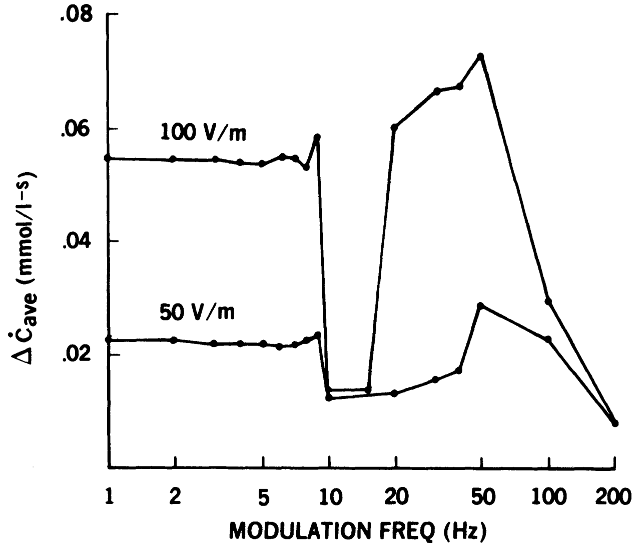


Fig. 2. Effect of change in modulation frequency, for field strengths 50 V/m and 100 V/m, on the arithmetic average differential.

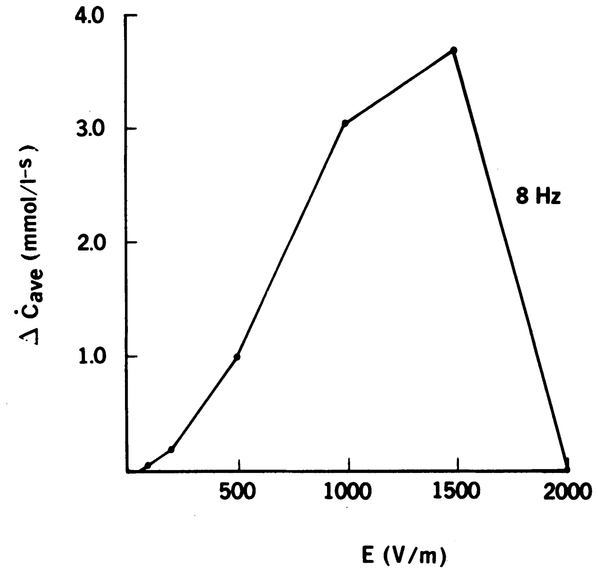


Fig. 4. Effect of change in field strength, for modulation frequency 8 Hz, on the arithmetic average differential.

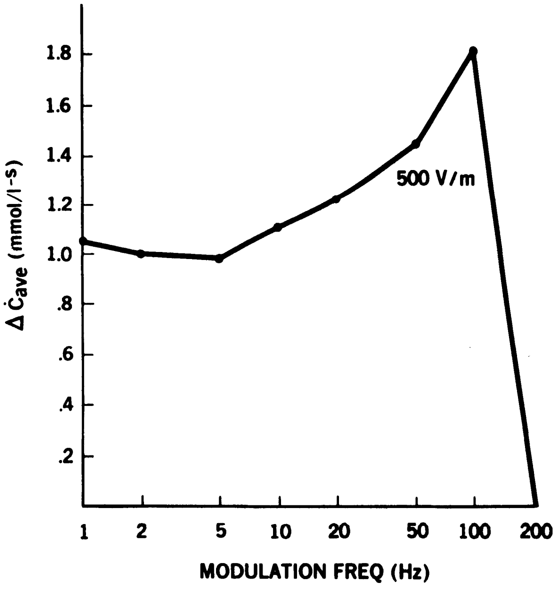


Fig. 3. Effect of change in modulation frequency, for field strength 500 V/m, on the arithmetic average differential.

4 shows that increasing field strengths can shut down the reaction. This occurs as the mean solvent temperature exceeds T_{max} of the enzyme temperature activity curve.

Insight into the sensitivity of product formation rate to modulation frequency is provided by analysis of the linearization of equations 3 and 4. The linearization is:

$$\dot{x} = -Ax - Bu$$

$$\dot{u} = Cx - Du + 1/2\ \sigma E^2 \sin^2(2\pi f t)$$

where

$$A = (h + k(T_{eq}))$$

$$B = \left.\frac{\partial k}{\partial T}\right|_{T_{eq}} \cdot X_{eq}$$

$$C \doteq \Delta H\ k(T_{eq})$$

$$D = \left(\ell - \Delta H \left.\frac{\partial k}{\partial T}\right|_{T_{eq}} \cdot X_{eq}\right)$$

$$x = X - X_{eq}$$

$$u = T - T_{eq}$$

X_{eq} = pre-irradiation equilibrium value of X

T_{eq} = pre-irradiation equilibrium value of T

The solutions of this linear system of equations are:

$$x = -\frac{B\sigma E^2}{4(AD+BC)} + \frac{B\sigma E^2}{4[(AD+BC-4\omega^2)^2 + 4\omega^2\ (A+D)^2]^{1/2}} \cos(2\omega t-\psi)$$

$$u = +\frac{A\sigma E^2}{4(AD+BC)} + \frac{(A^2+4\omega^2)^{1/2}\ \sigma E^2}{4[(AD+BC-4\omega^2)^2 + 4\omega^2\ (A+D)^2]^{1/2}} \sin(2\omega t-\psi-\theta)$$

$$\tan\psi = \frac{2\omega(A+D)}{(AD+BC-4\omega^2)}$$

$$\tan\theta = \frac{A}{2\omega}$$

$$\omega = 2\pi f$$

Notice that the driven oscillations in x and u are 180 degrees out of phase at low modulation frequencies but only 90 degrees out of phase at higher frequencies. This may partially explain the tendency for $\Delta\dot{C}_{ave}$ to generally increase with higher modulation frequencies until the high frequency cut-off point. However, also note that the amplitude of the driven oscillations in x and u have differing frequency dependencies, probably reflecting the two physical transport processes involved, namely, thermal and reactant diffusion. These differing frequency dependencies may be associated with the low frequency trough in the $\Delta\dot{C}_{ave}$ versus ω curves, at low field strengths particularly.

The work shown above addressing solvent heating can be directly extended to treat other possible effects. For example, by adding a term of the form

$$\alpha\cdot(E^2\sin^2\omega t).\ X$$

to the thermal balance equation, where α is an absorption constant, one can distinquish the effects of selective substrate heating. Additionally, letting various coefficients in the temperature-activity equation be field strength dependent, one can delineate potential direct field effects on chemical reaction mechanisms.

Thus, in summary, we feel that our work to date has brought forward significant evidence that chemical reaction systems can be perturbed by radiofrequency radiation in a manner which shows frequency sensitivity and complex responses to field amplitudes. This suggests the possibility that chemical reaction effects are the basis for reported EEG and calcium findings. We believe that chemical reaction-radiation interactions may be involved in the reported events for two reasons. First, the linearization analysis presented above predicts the existence of modulation frequency effects at very low field amplitudes, in fact, at amplitudes approaching zero volts/meter. Second, observe that as the temperature-activity curve approaches the steady state temperature of the unirradiated tissue, shut down of the reaction by increasing field strengths will occur at lower and lower field amplitudes. Finally, it also appears to us that non-ionizing radiation interactions with chemical systems are of great interest in themselves from the point of view of chemical theory and basic mechanisms of radiation action. We look forward to experimental work to test the application of the mathematical model to real systems.

APPENDIX

To derive equation 1, let A_n be the concentration of the active form of the enzyme, and A_d the concentration of the inactive or denatured form. At equilibrium, using the Arrhenius relation,

$$\kappa_1\exp(-\psi_1/T)A_n = \kappa_2\exp(-\psi_2/T)A_d$$

since $A_0 = A_n + A_d$,

$$A_n = \frac{A_0}{1 + (\kappa_1/\kappa_2)\exp[-(\psi_1 - \psi_2)/T]}$$

The rate of the chemical reaction X→C is $\kappa\exp(-\psi/T)A_n \cdot X = k(T)X$, so equation 1 follows.

Table I. System Parameters

h	=	Metabolite Diffusion Constant	= 3.75 sec^{-1}
ℓ	=	Thermal Diffusion Constant	= 150 sec^{-1}
ΔH	=	Heat of Reaction	= 1000 cal/mol
k_{max}	=	Maximum Value of Reaction Rate	= 13.86 sec^{-1}
T_{max}	=	Temperature at Maximum	= 313 °K
ψ	=	Activation Energy X→C	= 32,270 °K
β	=	Activation Energy $A_n \rightarrow A_d$	= 80,160 °K
σ	=	Tissue Conductivity	= 1.043 mho/m
T_e	=	External Temperature	= 310 °K
X_e	=	External Concentration	= 0.1 mol/liter

REFERENCES

Adey, W. Ross, 1979, Neurophysiologic Effects of Radiofrequency and Microwave Radiation, Bull. N.Y. Acad. Med. Vol 55, No 11, pg 163.

Bawin, S. M. et al., 1973, Effects of Modulated Very High Frequency Fields on Specific Brain Rhythms in Cats, Brain Res. Vol 58, pg 365.

Luben, Richard A. et al., 1982, Effects of Electromagnetic Stimuli on Bone and Bone Cells in Vitro: Inhibition of Responses to Parathyroid Hormone by Low-Energy Low-Frequency Fields, Proc. Natl. Acad. Sci. USA, Vol 79, pg 4180.

Servantie, B. et al., 1975, Synchronization of Cortical Neurons by a pulsed Microwave Field as Evidenced by Spectral Analysis of EEG from the White Rat, Ann. N.Y. Acad. Sci., Vol 247, pg 82.

Takashima, S. et al., 1979, Effects of Modulated RF Energy on the EEG of Mammalian Brains. Effects of Acute and Chronic Irradiation, Rad. and Environm. Biophys. Vol 16, 1979, pg 15.

ON THE GROWTH OF CELLULAR TISSUES UNDER CONSTANT AND FLUCTUATING ENVIRONMENTAL CONDITIONS

Rene Lefever
Chimie Physique II
University of Brussels
1050 Brussels, Belgium

Thomas Erneaux
Department of Engineering Science
and Applied Mathematics
Northwestern University
Evanston, Illinois 60201

INTRODUCTION

Today dissipative structures are at the center of interest of a noticeable part of the scientific community. This is particularly true in biology, where the discovery of reaction diffusion instabilities has opened a wholly new area of research on self organization. New types of dissipative structures, such as chemical clocks and reaction diffusion patterns, have become subjects of widespread interest. Their occurrence has been invoked in the explanation of a multitude of phenomena of self organization in biology. Examples include cellular differentiation, mitotic control, pattern formation, cellular communication, amoeba aggregation.[1] Generally speaking one may say that reaction diffusion instabilities show how molecules or cells can "communicate" over macroscopic times and distances using the very type of chemical processes and transport phenomena ubiquitously present in living systems.

What distinguishes dissipative structures from the organized states of matter studied by classical physics is that they are not governed by the usual (equilibrium) mechanisms of order-disorder transitions. In contrast to equilibrium structures, they are coherent behaviors which originate from irreversible transformations. Furthermore, only under far from equilibrium conditions do these transformations reveal their richness in ordered behaviors.

Dissipative structures are thus nonlinear phenomena which strongly depend on the properties of the environment. Two aspects are especially important in a biological context: (i) Many dissipative structures are frequently possible in a given system under given constant environmental conditions. Far from thermodynamic equilibrium, the behavior of nonlinear systems is generally not uniquely determined by the state of the environment, i.e. by the constraints acting upon the system. Very different "adaptative responses" may originate from small effects, e.g. for minute changes of the initial condition the evolution of a system may proceed towards quite different states of organization. (ii) The coupling between nonequilibrium systems and the environment is generally via multiplicative parameters, i.e. parameters whose value is fixed by the environment and which are multiplied by a function of the state of the system. When the environment fluctuates, the intensity of these fluctuations is modulated inside the system by the state function multiplying the external parameter. This makes the system highly sensitive to these fluctuations. When their variance increases, noise induced transition phenomena may occur.[2] These are sharp transitions which produce a qualitative change in properties and take place even if on the average the environment is strictly constant.

In this paper, we investigate the dynamics of cellular tissue growth. Obviously cellular replication is a nonlinear irreversible process strongly influenced by environmental factors. These features thus suggest that in a certain sense tissue growth must be assimilable to a behavior of a type of dissipative structure. Our aim is to give a precise content to this intuitive idea. For this purpose, we propose a reaction diffusion formulation of tissue growth. It permits us to associate this phenomenon with a class of coherent behaviors in the field of reaction diffusion instabilities. Our results illustrate for cellular tissues the aspects of far from equilibrium phenomena mentioned in (i-ii) above.

We apply our analysis more particularly to a specific problem, namely the growth of cancerous tissues attacked by immune cytotoxic cells. In this way, we situate the main outcomes of our approach within the context of an important biological problem. It is much beyond our capabilities to present a treatment which encompasses the overwhelming complexity of events at the in vivo level. Our starting point will be a kinetic model which describes the predator-prey relation of immune and cancerous cells as it can be observed in in vitro cytotoxic assays. The relevance of this model for the in vivo situation has been discussed elsewhere.[3 4] Throughout these notes our objective is thus to focus principally on the qualitative aspects of the results obtained.

The paper is organized as follows. First we model the phenomenon of tissue growth under constant environmental conditions. We show that it can be described by a reaction diffusion formalism and characterize some of the growth behaviors possible in the case of tumoral tissues. A multiplicity of regimes may occur for given environmental conditions. The choice between them is function of the initial condition. In a second part, we analyze the influence of environmental fluctuations on the system. A strong effect on the conditions for which the cancerous tissue is rejected is demonstrated.

TISSUE GROWTH IN A CONSTANT ENVIRONMENT

Reaction diffusion model

To begin with, let us present our reaction diffusion approach in a general manner. We shall particularize the problem in the next section where we study the competition between the growth of a tissue and immune cytotoxic reactions.

We consider a tissue constituted of cells existing in two physiological states: a state C in which they have the capability to replicate and a state D in which they have not. For the moment we do not specify the nature of these states more precisely nor the mechanism (reversible or irreversible) of the transition

$$C \rightleftarrows D \tag{1}$$

which transforms the cells from one type into the other. We assume that the spatial distribution D(r) of the D cells is given. r is the spatial coordinate; for simplicity the medium is one-dimensional. The first question we deal with is the derivation of a local balance equation continuous in time and in space which governs the evolution of the density of C cells.

Let N be the maximum number of cells which a spatial element of size Δr may contain and let a be the size of an individual cell. One has

$$N=\Delta r/a. \tag{2}$$

Suppose now that Δr's position on the spatial coordinate axis is r. At time t it contains C(r,t) cells of type C and D(r,t) cells of type D, $(C(r,t)+D(r,t)\leqslant N)$. We evaluate $C(r,t+\Delta t)$, where Δt is a short time interval with respect to the time scale λ^{-1} characterizing the replication of the C cells. Replication here is the overall process described in Fig. 1: a cell occupying in Δr a site i divides; the two daugh-

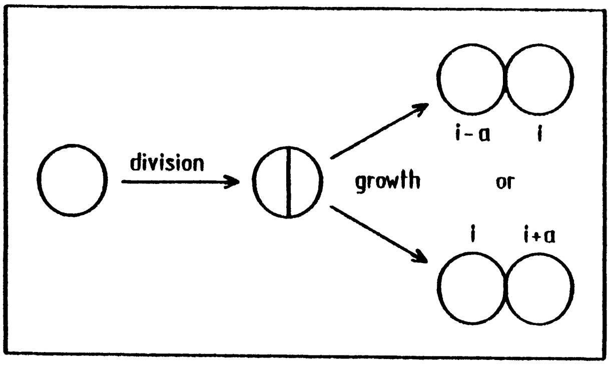


Figure 1: At replication a cell of size a located at a site i first divides and then the two daughter cells grow back to the original size a; they occupy the sites i and i+a or i-a. The relaxation time of the whole process is λ^{-1}

ter cells grow back to size a. One of these cells is then located at i and the other at i+a or i-a. The process can take place only if at least one of the sites at $i\pm a$ is empty. For $\Delta t \ll \lambda^{-1}$, we have that

$$
\begin{aligned}
C(r,t+\Delta t)=C(r,t)+\lambda[\,&C(r,t)\,.\frac{N-C(r,t)-D(r,t)}{N}-\\
&\frac{1}{2}\,\frac{C(r,t)}{N}\,.\,\frac{N-C(r+\Delta r,t)-D(r+\Delta r,t)}{N^2}-\\
&\frac{1}{2}\,\frac{C(r,t)}{N}\,.\,\frac{N-C(r-\Delta r,t)-D(r-\Delta r,t)}{N^2}+\\
&\frac{1}{2}\,\frac{C(r+\Delta r,t)+C(r-\Delta r,t)}{N}\,.\frac{N-C(r,t)-D(r,t)}{N^2}]\,\Delta t+o(\Delta t^2).
\end{aligned}
\tag{3}
$$

The first three terms between brackets come from the replication of cells belonging to Δr at t and giving rise to two daughter cells also belonging to Δr. The last term represents the replication of boundary cells belonging neighbouring volume elements at $r\pm\Delta r$ giving birth to a daughter cell in Δr. Introducing the densities

$$c(r,t)\equiv C(r,t)/N \text{ and } d(r,t)\equiv D(r,t)/N, \tag{4}$$

we take in (3) the limit $\Delta t\to 0$ and expand $c(r\pm\Delta r,t)$ and $d(r\pm\Delta r,t)$ in Taylor series around r. Taking (2) into account one obtains the desired equation:

$$\partial_t c=\lambda c(1-c-d)+\frac{\lambda a^2}{2}\,\partial_r[\,(1-d)\,\partial_r c+c\,\partial_r d]\,. \tag{5}$$

Let us remark that, though the cells are not mobile, (5) is a reaction diffusion equation. The cells invade space only by reproduction. The latter accounts both for the autocatalytic (logistic) term in (5) and for the nonlinear diffusion flux

$$J_c=-\frac{\lambda a^2}{2}\,[\,(1-d)\,\partial_r c+c\,\partial_r d]\,. \tag{6}$$

The first term in (6) expresses that replication can only take place in regions which are not totally occupied by D cells; the second term tells us that the spatial dissemination of the tissue is autocatalytic, i.e. it is modulated by a coefficient proportional to the density of cells capable to replicate. Equation (5) can easily be generalized to more than one spatial dimension. In that case, the existence of preferential directions of replication, due for example to the form of the cells, could be accounted for by making the replication constant function of the various spatial coordinates.

A simple particular case

In order to familiarize ourselves with (5), let us consider the simple situation where $d(r,t)=0$ everywhere. (5) then reduces to

$$\partial_t=\lambda c(1-c)+(\lambda a^2/2)\,\partial_{rr}c, \tag{7}$$

or

$$\partial_{t'}c=c(1-c)+\partial_{r'r'}c, \text{ with } t'=\lambda t \text{ and } r'= r/(a2^{-1/2}). \tag{8}$$

(8) has been introduced earlier by the geneticist R. A. Fisher[5] as a model for the spreading of a favorable gene in a population. It has been abundantly studied in the mathematical literature (for a review see ref. 6). Its properties relevant to the problem at hand here summarize as follows. (8) admits two uniform steady state solutions:

$$c_o=0 \text{ and } c_1=1. \tag{9}$$

c_o is unstable, while c_1 is stable according to linear stability analysis. Putting $c(-\infty,t)=c_1$ and $c(+\infty,t)=c_o$, (8) admits an infinity of travelling wave front solutions joining these two states. In first approximation these fronts are given by

$$c(\tau)=1/[1+\exp(-\tau/v)]^{-1}, \text{ with } \tau=r'-vt'. \tag{10}$$

v is the velocity with which the waves propagate over space without deformation. It is subjected to the condition

$$v \geqslant v_{min}=2. \tag{11}$$

These waves are stable with respect to bounded perturbations and are attractors for distinct sets of initial conditions.

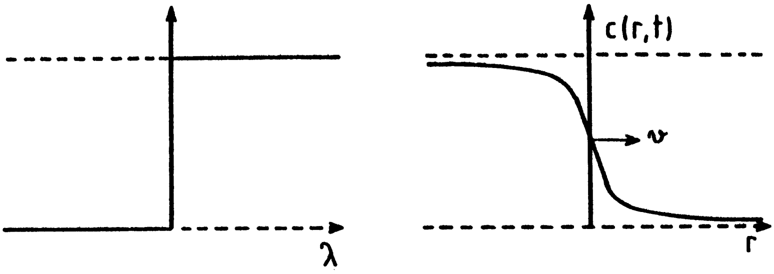


Figure 2: Bifurcation diagram of the uniform steady state solutions c_o and c_1. The steady states exchange stability at $\lambda=0$ (the unstable states are in broken line). For $\lambda>0$, a travelling wave front joining c_1 and c_o and propagating in favor of c_o is sketched.

When the supply of nutrients, oxygen etc. to the cells is constant, cellular tissues grow in an unbounded fashion. This unbounded growth behavior can be interpreted as the advance in space of a travelling wave of the type found in Fisher's equation (8). More exactly, the appropriate solution here is the one corresponding to $v_{min}= a\lambda$ (in the original time and space coordinates t and r). It agrees with the intuitively obvious guess that the motion of this simple tissue proceeds over space at the rate at which the tissue's boundary cells replicate; clearly this rate is equal to $a\lambda$. The biological relevance of v_{min} is further confirmed by the fact that the wave having this velocity is the attractor of all initial conditions of the form

$$c(r,0) = \begin{cases} 1 \text{ for } r<r_1 \\ \phi(r) \text{ for } r_1<r<r_2 \\ 0 \text{ for } r_2<r \end{cases} \qquad (12)$$

where r_1 and r_2 are finite and $\phi(r)$ is a monotonic and continuous function of r. Obviously all initial conditions biologically significant are of the form (12).

These results give confidence about the soundness of the approach. They suggest that the outcome of our analysis in the next section, which are hardly guessable by intuitive reasoning, are of predictive value.

ON THE GROWTH OF TUMORS ATTACKED BY THE IMMUNE SYSTEM

Kinetic equations for growth and cell-mediated cytolysis in tumors

The observations made so far show that the lysis of cancer cells (target cells) by immune T-lymphocytes (effector cells) can, broadly speaking, be described by the sequence of events:[7]

(i) Recognition step: a cell-cell contact forms between target and effector cell. This process is normally irreversible and involves a large number of receptors.

(ii) Progammation for lysis: the bound effector cell induces the unfolding of a "program" target cell whose conclusion is the lysis of the target cell (lethal hit).

(iii) Lysis step: once the lysis program has been triggered, the target cell carries it through irreversibly. This does not require the maintainance of the bond between target and effector cell.

Several authors[8,9,10] have proposed that the kinetics of (i-iii) can be represented by the two step reaction

$$E_o \xrightarrow[bC]{} E_1 \xrightarrow[\ell]{} E_o + D. \qquad (13)$$

The effector cell E_o first binds the target cell C at a rate bCE_o. The bicellular complex E_1 subsequently dissociates at a rate ℓE_1 into a target cell D programmed for lysis and an effector cell E_o ready again to participate to another cytolytic cycle. The similarity of this reaction scheme with the catalytic cycle of Michaelian enzymes is evident. This analogy accounts satisfactorilly of the kinetic behavior found with total (heterogeneous) lymphoid populations. Also other effector cells than T-lymphocytes follow a Michaelian type of law. Some representative values of the kinetic parameters introduced in (13) are given in Table I. Adopting for the growth process the model of the preceding section, the kinetic equations describing effector-

Table I. Approximated values of cytotoxic constants for the effector cells indicated (references to the original papers can be found in ref. 9).

Cytotoxic cell	b (day^{-1})	ℓ (day^{-1})
Allosensitized T-lymphocytes	18	18
Immune T-lymphocytes	0.43	0.85
Syngeneic natural killer cells	1-6	0.6-3
Syngeneic activated macrophages	0.1-0.4	0.2-0.7

target systems are:

$$\partial_t c=\lambda c[1-(c+d)]-bce_o+\lambda a^2\partial_r[(1-d)\partial_r c+c\partial_r d]/2 \tag{14}$$

$$\partial_t e_o=-bce_o+\ell e_1+D_o\partial_{rr}e_o \tag{15}$$

$$\partial_t e_1= bce_o-\ell e_1+D_1\partial_{rr}e_1 \tag{16}$$

$$\partial_t d=\ell e_1-md. \tag{17}$$

e_o and e_1 are the number of free and bound cytotoxic cells per unit volume. Cytotoxic cells reach the tumor by infiltration and their motion is in good approximation assimilable to a random walk. Hence, we describe it here by a Fickian law with diffusion coefficients D_o and D_1. b and ℓ are the constants of binding and lysis. (17) describes the elimination of the dead cancer cells; the rate of this process is simply supposed equal to md.

Behavior of uniform steady state solutions

We introduce the dimensionless parameters

$$\beta\equiv be_T/\lambda,\quad \theta\equiv\ell/b,\quad \gamma\equiv\ell\lambda/mb. \tag{18}$$

The uniform steady state solutions c_s of (14-17) are then given by the solutions of the equations

$$c_s[(1-\beta)\theta+(1-\theta-\beta\theta/\gamma)c_s-c_s^2]=0, \tag{18}$$

$$e_{os}=e_T/(1+c_s/\theta),\ e_{1s}=e_Tc_s/(1+c_s/\theta),\ d_s=\beta c_s/[\gamma\theta(1+c_s/\theta)] \tag{19}$$

with

$$e_T=e_{os}+e_{1s}.$$

The steady states c_s are sketched in Fig. 3. Their properties are

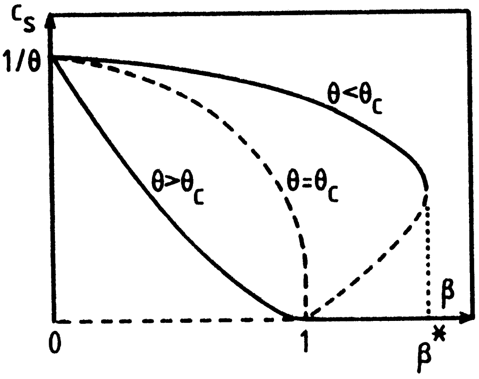


Figure 3: Sketch of the uniform steady states c_s as a function of β. Unstable states are represented in broken line.

summarized as follows:
(i) For

$$\theta>\theta_c\equiv\gamma/(1+\gamma) \tag{20}$$

(18-19) admit two steady states in the range $0\leqslant\beta<1$, namely $c_s=c_o=0$ which is unstable and a steady state corresponding to a finite cancerous population which is stable. $\beta=1$ is a bifurcation point where these two steady state branches meet each other. For $\beta>1$ only the state c_o is physically acceptable.
(ii) The point

$$\theta=\theta_c,\ \beta=\beta_c=1 \text{ and } c_s=c_c=0 \tag{21}$$

is a critical point at which the slope $dc_s/d\beta$ of the cancerous branch of steady states is infinite.
(iii) For $\theta<\theta_c$, the steady state curves exhibit a region of bistability. In the range

$$1\leqslant\beta\leqslant\beta^*, \tag{22}$$

(18) admits two stable steady state solutions, c_o and the upper branch of cancerous states; the steady state existing in between is unstable. For $\beta<1$, the behavior is qualitatively similar to (i).

In the remaining part of this section, we shall analyze the properties of the traveling waves joining c_o and the cancerous branch of steady states.

Travelling wave solutions of (14-17)

The mathematical problem of finding travelling wave front solutions of reaction diffusion equations has been studied in great detail using phase plane techniques for one variable systems[6] (see ref. 11 for an application to a simplified version of the present model). Two variable models have been studied singular perturbation methods.[12,1] These techniques are difficult to use when the number of variables is larger than two. Recently however we have proposed a perturbation method permitting to construct slowly-varying wave front solutions valid for many variable systems.[14] This method makes no assumptions on the time scales of evolution of the variables. It is well adapted to conditions where the values of the two uniform steady state solutions joined by the waves are close to each other. Looking at Fig. 3, these conditions are typically met by our model in the neighbourhood of the critical point (21). We sketch here the results obtained in that case (for a more detailed mathematical derivation see ref. 14).

We introduce the dimentionless parameters

$$\rho \equiv \ell/\lambda, \quad \mu \equiv D_o/(\lambda a^2), \quad \nu \equiv D_1/(\lambda a^2), \quad t' \equiv \lambda t \tag{23}$$

and redefine the variables as follows

$$x \equiv bc/\ell K, \quad y \equiv (1+K)(1-e_o)/K, \quad w \equiv (1+K)(1-e_o-e_1)/K, \quad p \equiv \gamma bd(1+K)/\beta\ell K \tag{24}$$

where $K \equiv bc_s/\ell$ corresponds to the stable cancerous steady state. Equations (14-17) may now be rewritten as

$$\partial_{t'} u = F(u) + \partial_{r'}[D(u)\partial_{r'} u] \tag{24}$$

where $u \equiv (x\ y\ w\ p)$ is the vector of the variables,

$$F(u) \equiv \begin{pmatrix} (1-\beta)x - \theta Kx^2 + \beta Kxy/(K+1) - \beta\theta Kxp/\gamma(K+1) \\ \rho[(K+1)x - y + w - Kxy] \\ 0 \\ \gamma[(K+1)x - p - Kxy] \end{pmatrix} \tag{25}$$

and

$$D(u) \equiv \begin{pmatrix} 1-\beta\theta Kp/\gamma(K+1) & 0 & 0 & \beta\theta Kx/\gamma(K+1) \\ 0 & \mu & 0 & 0 \\ 0 & (\mu-\nu) & \nu & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{26}$$

In the new variables, the steady state corresponding to the cancerous state is given by

$$(x_s\ y_s\ w_s\ p_s) = (1, 1, 1, 1), \text{ for all } \beta.$$

To apply the perturbation method, we must work in the neighbourhood of the critical point, i.e. where the value of K is small. We define the smallness parameter ε by

$$I\varepsilon \equiv \theta_c - \theta \text{ with } \begin{cases} I=1 \text{ for } \theta<\theta_c \\ I=-1 \text{ for } \theta>\theta_c, \end{cases} \tag{27}$$

and expand β and K as

$$\beta=1+\varepsilon^2 b_2, \quad K=\varepsilon K_1+\varepsilon^2 K_2+\ldots \tag{28}$$

At the lowest order, one has

$$\begin{aligned} K&=I\varepsilon[\,1+I(1-4b_2\theta_c^3)^{1/2}]\,(2\theta_c^2) \\ &= K_1 \end{aligned} \tag{29}$$

with

$$b_2 \in [\,-\theta_c/4,0] \text{ for } I=-1 \text{ and } b_2 \in [\,-\theta_c/4,\theta_c/4] \text{ for } I=1. \tag{30}$$

Travelling wave front solutions of (24) are stationary state concentration profiles in a reference frame moving at the same velocity v as the front, i.e. they are solutions of the equation

$$d_\tau[\,D(u)d_\tau u]+vd_\tau u+F(u)=0, \text{ where } u(\tau)=u(r'-vt'). \tag{31}$$

The boundary conditions are

$$u(-\infty)=(1\ 1\ 1\ 1)\equiv u(1), \quad u(+\infty)=(0\ 0\ 0\ 0)\equiv u(0). \tag{32}$$

We now consider successively the case of systems working above and below the critical point (21).

$\theta>\theta_c$ and $I=-1$. By linear stability analysis one easily verifies that u(0) is an unstable focus unless

$$v \geqslant v_{min} = (-b_2)^{1/2}. \tag{33}$$

This condition must be satisfied for the solutions to be physically acceptable and is equivalent to the condition found in the scalar case (11). Similarly to the scalar case, the system admits for $\theta>\theta_c$ an infinity of travelling wave front solutions. To see this, we introduce

$$\xi \equiv \varepsilon^2 \tau \tag{34}$$

and expand $u(\xi)$ as

$$u(\xi)=u_o(\xi)+\varepsilon u_1(\xi)+\ldots \tag{35}$$

Replacing u in(31), and going to order ε^0 we must satisfy the equation

$$F_o u_o = 0, \tag{36}$$

where

$$F_o \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\rho & \rho & -\rho & 0 \\ 0 & 0 & 0 & 0 \\ -\gamma & 0 & 0 & \gamma \end{pmatrix}. \tag{37}$$

The solution of (36) is

$$u_o(\xi) = (a(\xi)\ \ b(\xi)\ \ [-a(\xi)+b(\xi)]\ \ a(\xi))^{tr}. \tag{38}$$

The functions a and b remain undetermined at this order. At the order ε^1, one finds that

$$F_o u_1 = \begin{pmatrix} -K_1 a^2 + K_1 ab \\ K_1 a(1-b) \\ 0 \\ K_1 a(1-b) \end{pmatrix} \equiv F_1(a,b). \tag{39}$$

$|F_o| = 0$, implying that the solvability condition

$$F_o^+ u_o^+ = 0, \quad \langle u_o^+, F_1(a,b)\rangle = 0 \tag{40}$$

must be fulfilled, where F_o^+ and u_o^+ are the matrix adjoint to F_o and its eigenvector. (40) yields that

$$a(\xi) = b(\xi). \tag{41}$$

Furthermore, it is easy to see from the equation for w at the order ε^3 that

$$w_1(\xi) = 0. \tag{42}$$

Taking (41-42) into account the solution of (39) is

$$u_1(\xi) = (e(\xi)\ \ \ f(\xi)\ \ \ 0\ \ \ [e(\xi) + K_1 a(\xi)(1-a(\xi))])^{tr} \tag{43}$$

where the functions e(ξ) and f(ξ) are undetermined. At this stage, it may seem that we have worsened our lot since two new undetermined functions e,f have arised. At each order ε^k however, we must satisfy a solvability condition of the form

$$\langle u_o^+, F_k(u_{k-1}(\xi), \ldots, u_o(\xi))\rangle = 0, \tag{44}$$

where F_k is a vector whose components are functions of the solutions at the lower orders in ε. Proceeding to the order ε^2, (43), with k=2, yields after some manipulation a closed form equation for a(ξ), namely

$$v d_\xi a = a(1-a)(b_2 - \theta_c K_1^2 a). \tag{45}$$

Solving (45) for $a(\xi)$, the solution of (36) is now completely determined. It reads

$$u_o(\xi) = a(\xi)(1 \quad 1 \quad 0 \quad 1)^{tr} \tag{46}$$

with $a(\xi)$ given by

$$\frac{a^{1/q}(1-a)^{1/(q-1)}}{(a-q)^{1/[q(q-1)]}} = \exp[\theta_c K_1^2 \xi/v], \text{ and } q \equiv b_2/\theta_c K_1^2. \tag{47}$$

Clearly (46-47) admit an infinity of travelling wave front solutions for $v \geqslant v_{min}$ joining the steady states u(0) and u(1).

$\theta<\theta_c$, $I=1$ and $b_2 \in (0, \theta_c/4]$. The main difference with the preceding case is that now both u(1) and u(0) are stable. The wave fronts joining these two states can be obtained using the perturbation method used above. We shall therefore simply report the results and skip the detailed mathematical derivation which will be presented elsewhere (14). Two distinct classes of wave solutions are found.

The first class is obtained putting

$$\xi \equiv \varepsilon \tau \tag{48}$$

and by expanding $u(\xi)$ and β as in (28-35). Furthermore, since both boundaries are stable steady state solutions, the conditions on the velocity are more stringent. It can no longer be left as an arbitrary parameter. Expanding it as

$$v = \varepsilon(v_o + \varepsilon v_1 + \ldots), \tag{49}$$

one finds that at the lowest order in ε the solution satisfying the boundary conditions (32) is given by*

$$u_o(\xi) = a(\xi)(1 \ 1 \ 0 \ 1)^{tr} \tag{50}$$

with

$$a(\xi) = 1/[1 + \exp(\theta_c^{1/2} K_1 \xi/\sqrt{2})] \tag{51}$$

and

$$v_o = K_1(\theta_c/2)^{1/2}(1 - 2b_2/\theta_c K_1^2). \tag{52}$$

---

*Though (51-52) do not depend explicitely on the diffusion coefficients, the latter are not arbitrary. At the order ε^2, the solvability condition has been fulfilled by putting $\mu=\nu$.

For $v_o>0$, the tumor is progressive. Though locally the normal state $u(0)$ is stable, globally it is metastable with respec to the tumoral state $u(1)$. At

$$\beta=\beta_{coex}\equiv1+2\varepsilon^2/9\theta_c^3<\beta^*=1+\varepsilon^2/4\theta_c^3, \qquad (53)$$

$v_o=0$: both uniform steady states are equally stable; β_{coex} is analogous to an equilibrium point of phase coexistence. Beyond this point the tumor regresses.

The second class of wave solutions is obtained putting

$$\xi\equiv\varepsilon^{3/2}\tau$$

and expanding v as

$$v=\varepsilon^{1/2}(v_o+\varepsilon v_1+\ldots).$$

At the lowest order in ε, the solution satisfying the boundary conditions (32) is given by

$$u_o(\varepsilon)=a(\varepsilon)(1,1,0,1)^{tr} \qquad (54)$$

with

$$a(\varepsilon)=1/\ 1+\exp[\ (b_2(\mu-\nu)/\theta_cK_1)^{1/2}\xi] \qquad (55)$$

and

$$v_o=-(b_2(\mu-\nu)/\theta_cK_1)^{1/2}. \qquad (56)$$

Clearly this solution only exists for $\mu>\nu$. This condition however is no limitation in the case of the present model. It is indeed plausible that the mobility of the bound cytotoxic cells is lower than that of the free ones. In contrast with the preceding class of travelling wave front solutions, v_o does not change sign in the interval $1<\beta<\beta^*$. The point of coexistence β_{coex} coincides with $\beta=1$. In the whole bistability region, the tumoral wave is regressive.

We shall comment more in detail on the properties and biological signification of the travelling wave front solutions reported here to describe the growth of tumoral tissues in the last section. Before doing so, we would like to investigate, in the case of cellular tissues, the other important aspect of nonequilibrium nonlinear systems, namely their great sensitivity to environmental fluctuations. We consider this problem in the next section and show especially that the occurrence of bistability strongly depends on these fluctuations.

INFLUENCE OF ENVIRONMENTAL VARIABILITY

We have seen that the properties of the travelling waves describing cellular tissues change completely when one passes from a region in parameter space where the uniform steady state solutions present bistability to a region where they do not (cf. Fig. 3). In order to understand the mechanism of evolution of a tissue it is thus essential to know whether or not this behavior is involved. In the presence of environmental fluctuations the answer to that question may turn out to be more complex than usually thought.

Up to now we have been content with the idealization of an environment which is strictly constant in time. The results of this section show the limitations to which this approach may be subjected, particularly in the case of natural systems such as cellular tissues. Cellular environments depend on a multitude of factors (temperature, radiations, supply of oxygen and nutrients, hormones, etc.) which to a greater or lesser degree fluctuate in the course of time. Through cumulative effects of these fluctuations of the environment, many cellular parameters in turn become fluctuating quantities. At best they are constant on the average, and exhibit a time behavior with an appreciable amount of external noise. The analysis of the tumor growth model in the presence of noise of this type yields results whose main features are that (i) the stationary state behavior remains remarkably coherent even in the presence of a large, extremely rapid and completely incoherent (memoryless) noise like gaussian white noise; (ii) by increasing the variance of this noise, it is possible to induce bistability in a system which displays none under constant environmental conditions and to modify in this way the mechanism of tumor growth. Rejection of the tumor seems facilitated.

We shall be interested in fluctuations affecting the binding constant b. In in vivo tumoral systems this parameter generally is a fluctuating quantity. Indeed the populations of immune T-lymphocytes found in tumors are heterogeneous. They are constituted of several clones which recognize the target's antigens with different degrees of specificity.

To reduce the discussion to the essential, let us consider the case where the dead cancer cells D are eliminated rapidely so that their population remains small and needs not to be taken into account into the equations ($\gamma \to \infty$). Furthermore, we may drop the diffusion terms in the kinetic equations; because we want to investigate the behavior of the uniform steady states of these equations. We suppose that the fluctuations of b take place around a fixed average value. We write b(t) in the form

$$b(t)=b+\lambda z(t), \text{ with } \langle b(t)\rangle=b \text{ and } \langle z(t)\rangle=0. \tag{57}$$

z(t) is an Ornstein-Uhlenbeck noise whose temporal evolution is given by the Langevin equation

$$d_t z = -\frac{2}{\tau_{cor}} z + \bar{\sigma}\xi(t). \tag{58}$$

τ_{cor}, $\bar{\sigma}$ are the correlation time and the variance of z; $\xi(t)$ is the standard gaussian white noise. Taking (57) into account, the kinetic equations now read

$$d_t c = \lambda c(1-c) - bce_o - \lambda zce_o \tag{59}$$

$$d_t e_o = \ell(e_1 - bce_o/\ell - \lambda zce_o/\ell), \text{ with } e_T = e_o + e_1. \tag{60}$$

We define

$$x \equiv bc/\ell, \quad y \equiv e_o/e_T, \quad \tau \equiv \lambda t \tag{61}$$

and introduce the parameters β, θ and ρ defined in (22-28). Equations (58-60) become

$$d_\tau z = -\; 2z/(\lambda\tau_{cor} + \bar{\sigma}\xi(\tau)/\lambda^{1/2} \tag{62}$$

$$d_\tau x = x(1-\theta x) - \beta xy - ze_T xy \tag{63}$$

$$d_\tau y = \rho[\,1-(1+x)y - \lambda zxy/b]\,. \tag{64}$$

We take as smallness parameter

$$\varepsilon \equiv \lambda\tau_{cor}/2 = \tau_{cor}/2\tau_{macro} \tag{65}$$

which espresses the ratio by which the time scale of the evolution of external noise is much faster than the time scale of evolution of x (given in first approximation by $\lambda^{-1} = \tau_{macro}$). The total number of cytotoxic cells per unit volume is large ($e_T = 10^3 - 10^4$ cells/mm^3). We put

$$e_T = e/\varepsilon, \text{ where } e = o(1) \text{ take } e=1 \text{ in the following.} \tag{66}$$

With these scalings (62-64) can be rewritten as

$$d_\tau x = x(1-\theta x) - \beta xy - zxy/\varepsilon \equiv f(x,y) + zg(x,y)/\varepsilon \tag{67}$$

$$d_\tau y = \rho[\,1-(1+x)y - \varepsilon^2 2zxy/b\tau_{cor}] \equiv \rho[\,h(x,y) - \varepsilon^2 2zxy/b\tau_{cor}] \tag{68}$$

$$d_\tau z = -\; z/\varepsilon^2 + \sigma\xi(\tau)/\varepsilon, \text{ where } \sigma \equiv \bar{\sigma}^2/2\gamma. \tag{69}$$

The triple (x,y,z) is a markovian diffusion process; accordingly the temporal evolution of its transition probability density $p(x,y,z|0)$ is given by the Fokker-Planck equation

$$\partial_\tau p(x,y,z|0) = \{-\partial_x f(x,y) - (z/\varepsilon)\partial_x g(x,y) + (1/\varepsilon^2)(\partial_z z + (\sigma^2/2)\partial_{zz}) - \rho[\,\partial_y h(x,y)] + \varepsilon^2\rho(2\gamma/b)z\partial_y xy\} p(x,y,z|0). \tag{70}$$

We are interested in the stationary state behavior of (70), more precisely in the behavior of the stationary probability density $p_s(x)$. Its extrema are the stochastic equivalents of the uniform steady states in the deterministic description; the maxima correspond to stable states, the minima to unstable ones (for a thorough justification of this interpretation, see ref. 15). When the external noise vanishes, i.e. $\sigma=0$, the location of the extrema of $p_s(x)$ coincides exactly with the value of the uniform steady states. For $\sigma\neq 0$, this is no longer the case. The influence of the noise z(t) in (67) is modulated by the state function xy; it is multiplicative. As we show below, this gives rise to an interplay between the noise and the nonlinearities of the system which deeply modifies the bifurcation diagram of Fig. 3.

It is not possible to obtain the general stationary state solution of (70) for any value of σ and γ. The results presented below are obtained in the white noise limit, i.e. for $\varepsilon\to 0$. Furthermore, we assume that the cytolytic process is rapid compared to the rate of cellular replication: $\rho>>1$. Under those conditions, it is possible, using the wide band perturbation method[15,16,17] to find an approxite solution to the problem. Depending on the rate of cytolysis with respect to the rate of the noise one finds the following (the mathematical derivation will be presented in (18), see also (19)).

(i) $\varepsilon^{-2}>>\rho$: external noise is faster than cytolysis. Then

$$p_s(x)=\frac{N(1+x)^2}{x^{[1-\frac{2(1-\beta)}{\sigma^2}]}}\exp\frac{2}{\sigma^2}[(2-\theta-\beta)x+(1/2-\theta)x^2-\theta x^3/3] \tag{71}$$

(ii) $\rho>>\varepsilon^{-2}$: cytolysis is faster than external noise. Then

$$p_s(x)=\frac{N(1+x)}{x^{[1-2\frac{(1-\beta)}{\sigma^2}]}}\exp\frac{2}{\sigma^2}[(2-\theta-\beta)x+(1/2-\theta)x^2-\theta x^3/3]. \tag{72}$$

$N^{-1}\equiv\int_0^\infty p_s(x)dx$ is a normalisation constant.

Figure 4 illustrates the behavior of $p_s(x)$ in (71) and (72), respectively in (a-a') and (b-b'). The extrema of $p_s(x)$ corresponding to finite values of x, i.e. of the cancerous population, are plotted in term of the average environmental state (given by β) and for different values of the noise intensity (given by σ). For $\sigma^2=0$, we recover the deterministic curve given by (23) with $\theta=1,\gamma^{-1}=0$. It is a monotonously decreasing function of β. In the presence of external noise, bistability appers. When σ^2 increases, the bistability region becomes larger and moves towards the lower values of β. The random variations of the binding constant b(t) facilitate the rejection of the tumor by the immune system. As a whole, the efficiency of a heterogeneous cytotoxic population whose binding

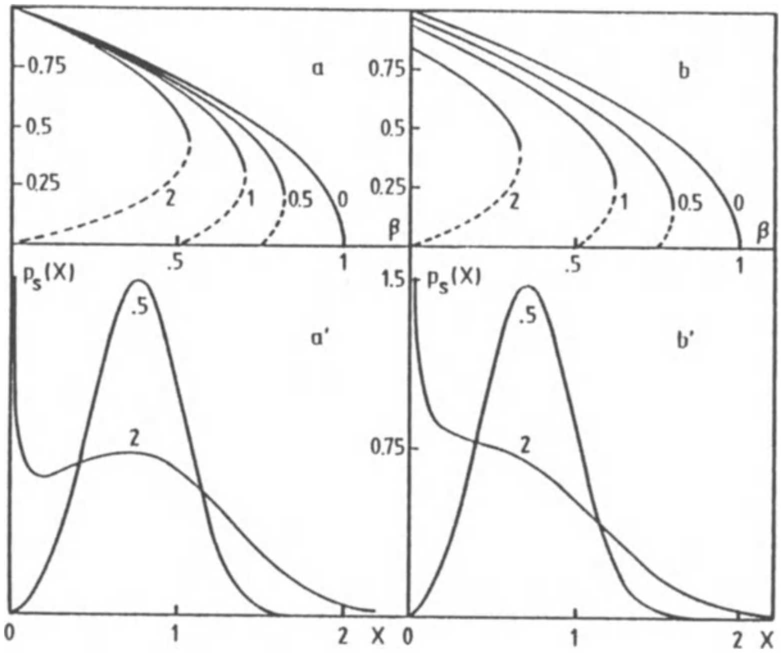


Figure 4: In (a, b) the extrema of $p_s(x)$ are plotted as a function of the average value of β and for the values of σ^2 indicated; 0 = 1. The maxima are represented in full line and the minima in broken line. In a', b') $p_s(x)$ as given by (71) and (72) is represented for $\beta = 0.5$ and two values of σ^2. Extinction of the x population is favored for $\sigma^2 = 2$ as compared to $\sigma^2 = 0.5$. The maxima which persists in a fluctuating environment disappears more rapidly in the case where σ^2 increases.

constant b(t) is gaussian distributed around some average value, is greater than the efficiency of the homogeneous population which would have a fixed binding constant equal to this average. Remarkably also, the behavior of the system remains quite organized even for large values of σ^2. The probability density is not simply smeared out by the noise, but exhibits sharply defined transition phenomena. Indeed, two well-defined regimes, corresponding to completely different shapes of $p_s(x)$ (see Fig. 4), are separated by a sharp threshold value of σ. For small values of x, $p_s(x)$ in (71-72) behaves like

$$p_s(x) \simeq 1/[x^{1-2(1-\beta)/\sigma^2}].$$

Accordingly, for $\sigma^2/2<1-\beta$, $p_s(x)\to 0$ as $x\to 0$. The only extremum of $p_s(x)$ is the maximum which is the continuation of the deterministic state. At $\sigma^2/2=1-\beta$, the shape of the probability density changes abruptly. For $1-\sigma^2/2<\beta<1$, it is still normalizable, but it diverges at x=0. One is in the bistability region where rejection becomes possible. For $\beta>1$, $p_s(x)$ is a δ-peak on $x=0$. Extinction is certain.

CONCLUSIONS

There is no doubt that cellular tissues exchange energy and matter with their environment and hence are dissipative systems whose evolution is, thermodynamically speaking, a nonequilibrium phenomenon. In these notes, we have tried to go beyond this elementary evidence and to propose a precise physical description of these systems. We have shown that the evolution of tissues can be situated in the main stream of the self organization phenomena observed in reaction-diffusion systems. The transitions and nonequilibrium behaviors which tissues display are typical of far from thermodynamic equilibrium conditions. In this sense, it is justified to call their spatio-temporal organization, as modelled by travelling wave fronts, dissipative structures. It is clear that the in vivo situations are more complex than the models which we have presented and of which eq. (5) is the core. We believe however that the assumptions made in deriving this equation correspond to really basic facts and thus that at least in part, the properties related to these facts play a role in the mechanisms of progression and regression of tissues.

An advantage of the approach proposed is that it establishes a direct connection between the properties of the uniform tissue "phases" and the overall space time behavior. These phases are function of cellular parameters (λ, b, ℓ, diffusion coefficients, etc.) which are accessible independently by in vitro measurements. The results therefore should be amenable to experimental testing. The existence of a multiplicity of uniform steady state branches and of transition points where these branches exchange stability (cf. Fig. 2-3) evidently is a consequence of the nonlinearity and of the irreversibility of the processes involved (replication, cytolysis). Judging from the numerical values of the constants reported in Table I, it is likely that a bistability phenomenon is controlling the growth of certain tumoral systems; the more so when the level of external noise affecting parameters such as b becomes important. For example, when the elimination of dead cancer cells is rapid, the critical value θ_c is close to one. Given that $\theta=b/\ell$, many effector-target systems seem to function in the neighbourhood of θ_c. In summary our model of effector-target interactions mainly indicates the following. In the course of the immune response against tumors, two distinct situations may arise in regard to the conditions of tumor rejection. The first situation corresponds to $\theta>\theta_c$. In that case the tumor is progressive as long as $\beta<1$. The travelling wave describing this progression joins at $r=-\infty$ and $r=+\infty$ the stable state u(1) to the unstable state u(0). Qualitatively the behavior is identical to the scalar case (8). Notably the velocity of propagation of the wave does not depend in first approximation on the mobility of the cytotoxic cells, i.e. on their diffusion coefficient. The second case arises for $\theta<\theta_c$. Two mechanisms of growth are then possible, corresponding to the waves given by (51-52) and (55-56). Notably these waves differ by the location of the rejection transition point given by β_{coex} and by their dependency on diffusion coefficients.

The results for b fluctuating are examplative of the influence of complex environmental conditions. In particular, processes sensitive to radiations could lead to noise induced transitions: absorption indeed typically is a multiplicative stochastic process when the intensity fluctuates. The results offer a glimse into a domain which is still underinvestigated theoretically and experimentally. There exists some gut feeling against the idea that a form of environmental variability assimilable to a noise may play a significant role and be a source of coherent behaviors. This however may only be a prejudice stemming from our desire to reduce the complexity of our descriptions and our habit to think about selforganization in strictly deterministic terms.

ACKNOWLEDGMENT

This work is part of a research program financially supported by the National Foundation for Cancer Research.

REFERENCES

1. For a general presentation of the field and references to original papers see: G. Nicolis and I. Prigogine, Selforganization in nonequilibrium systems, Wiley, New York (1977).
2. See the papers by Horsthemke, Lefever, Sancho and San Miguel in: Stochastic nonlinear systems, Springer Series in synergetics 8 (1981).
3. R. Garay and R. Lefever, J. Theor. Biol. 73, 417 (1978).
4. R. Lefever and R. Garay, in: Biomathematics and cell kinetics, Elsevier, Amsterdam (1978)
5. R. A. Fisher, Ann. of Eugenics 7, 355 (1937)
6. P. C. Fife, Mathematical aspects of reacting and diffusing systems, Springer-Verlag, Heidelberg (1979).
8. R. M. Thorn and C. S. Henney, J. Immunology 117, 2213 (1976).
7. C. S. Henney, C. R. C. Crit. Rev. Immunol.
9. R. Garay and R. Lefever, C. R. Acad. Sci. Paris 285D, 741 (1977).
10. S. J. Merril, Math. Biosc. 62, 219 (1982).
11. I. Prigogine and R. Lefever, Comp. Biochem. Physiol. 67B, 389 (1980).
12. J. Rinzel and J. B. Keller, Biophys. J. 13, 1313 (1973).
13. P. Ortoleva and J. Ross, J. Chem. Phys. 63, 3398 (1975).
14. T. Erneux and R. Lefever, in Preparation.
15. W. Horsthemke and R. Lefever, Noise induced transitions, to appear, Springer-Verlag (1983).
16. G.Blankenship and G. C. Papanicolaou, SIAM J. Appl. Math. 34, 437(1978)
17. W. Horsthemke and R. Lefever, Z. Phys. 40, 241 (1980).
18. W. Horsthemke and R. Lefever, in preparation.
19. R. Lefever and W. Horsthemke, Bull. Math. Biol. 41, 469 (1979).

IONIC CURRENTS AND FIELD EFFECTS IN NEURAL EXTRACELLULAR SPACES

T. Triffet[1] and H.S. Green[2]

[1]College of Engineering and Theoretical Physics Division
University of Arizona Research Laboratories

[2]Department of Mathematical Physics
University of Adelaide, S. Australia

INTRODUCTION

A distinctive feature of animal life is the electrical activity displayed by individual cells as a result of the relatively large potential between the intracellular and extracellular electric fluids. This potential difference is established as a significant byproduct of metabolic activity within the cell, which reduces the concentrations of sodium and calcium ions internally to levels far below those in the extracellular environment. The electrical activity is associated primarily with the nervous system, and especially with the internal surface of the neural membrane where the largest variations of the electrical field are observed during the passage of action potentials. Of equal if not greater significance, however, are the graded potentials within the cells, and what we shall call "external potentials" in the extracellular fluid. The latter are detectable in an attenuated form as electric and magnetic fields outside the body, notably on the scalp where they form the basis of clinical and diagnostic studies of cortical activity by means of electroencephalograms (EEGs) and magnetoencephalograms (MEGs).

The records obtained in this way reflect the variations in potential near the surfaces of a large number of neurons and other cells in the neighboring cortex. The same variations can be detected, with amplitudes up to several millivolts, by suitably designed microelectrodes introduced into the extracellular fluid of the cortex. Some of the variations of potential observed in the extracellular fluid take the form of spikes or spindles, which are clearly associated with action potentials in individual neurons or synchronous potentials in related groups of neurons. In some instances the latter have a periodic character, especially the alpha-spindles (1) which are almost certainly synchronized by activity in specific nuclei of the thalamus. Other variations have a less localized source, such as those investigated by Freeman and his associates (2), and those associated with the alpha- and beta-rhythms which appear characteristically in waking states of relaxation. Under some circumstances the alpha-waves may be associated with the alpha-spindles, but they normally occur independently, and their source has been found by Lopes da Silva and Storm van Leeuwen (3) and Basar (4) in a layer rich in apical dendrites, about 1 mm below the cortical surface in the dog. We shall in the following elucidate the mechanism of these alpha- and beta-waves as well as of other types of extracellular potentials.

There is already a considerable body of knowledge concerning extracellular fields. One of the earliest detailed experimental studies was made by Lorente de Nó (5), who charted the external field associated with action potentials in a neuron of the bullfrog, though in an environment removed from the animal. He was also able to model the observed field satisfactorily by an analysis based on Helmholz's demonstration that the field in a region free of sources is completely determined by the values of the potential and its normal derivative on any closed surface containing the sources.

This type of analysis has been subsequently developed and extended by Plonsey (6, 7), and applied to determine extracellular electric and magnetic fields of a neuron in a medium with specified characteristics. His work to some extent complements that of Rall (8), who modeled the development of graded potentials in the dendritic trees which are a common feature of regions near the surface of the cortex. However, the potentials recorded *in situ* by introducing microelectrodes into the actual extracellular environment of the cortex. Towe and Freeman (9, 2) display features which are not satisfactorily accounted for by these theories, such as a reversal of polarity with distance from an active neuron, as shown in Figure 1. The explanation is evidently related to the special nature of the environment.

It is usual to make a distinction between the largely unmyelinated grey matter near the surface of the cortex and the myelinated white matter underneath. The grey matter is itself organized in two layers, the uppermost principally containing cells with an inhibiting action and a profusion of dendritic branches, such as the Purkinje cells. A little below them are the smaller cells of the granular layer; these are excitatory in action, but also have complex dendritic branches. There are no dendrodendritic

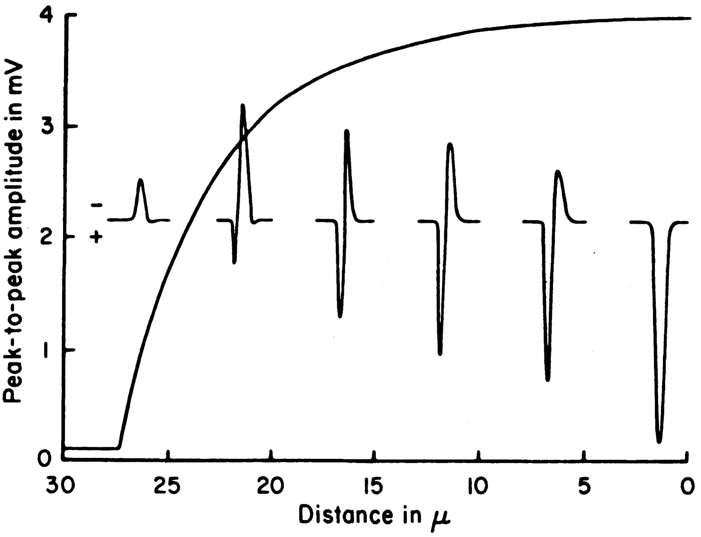


Figure 1. Growth in amplitude and changes in configuration of extracellular spike discharge as a microelectrode approaches the neural membrane (9).

or dendrosomatic synapses, which shows that the dendrites function almost exclusively as receivers of electrical impulses and the axons as transmitters, so far as the neural network is concerned (Figure 2). Nevertheless, axons and dendrites occupy only a small fraction of the total extracellular space; the remainder is occupied by glial cells. These are classified as oligodendrocytes, which associate themselves closely with myelinated fibers, and astrocytes, which predominate in the grey matter (Figure 3).

The layer of extracellular fluid separating neurons from glial cells has a remarkably uniform thickness of about 150 Å. It has been variously suggested that the function of these glial cells is to separate or insulate neurons, and that they may serve to regulate potassium concentrations in the extracellular fluid, either directly or by diffusion of water thorugh the membrane; but it remains true today, as in earlier times, that their essential role is a matter of conjecture. At one time they were regarded as electrically inactive, but there is now a very large body of evidence to the contrary (12). Since the glial cells have no axon, they do not exhibit action potentials in the ordinary sense, but they do have a membrane potential comparable with that of neurons,

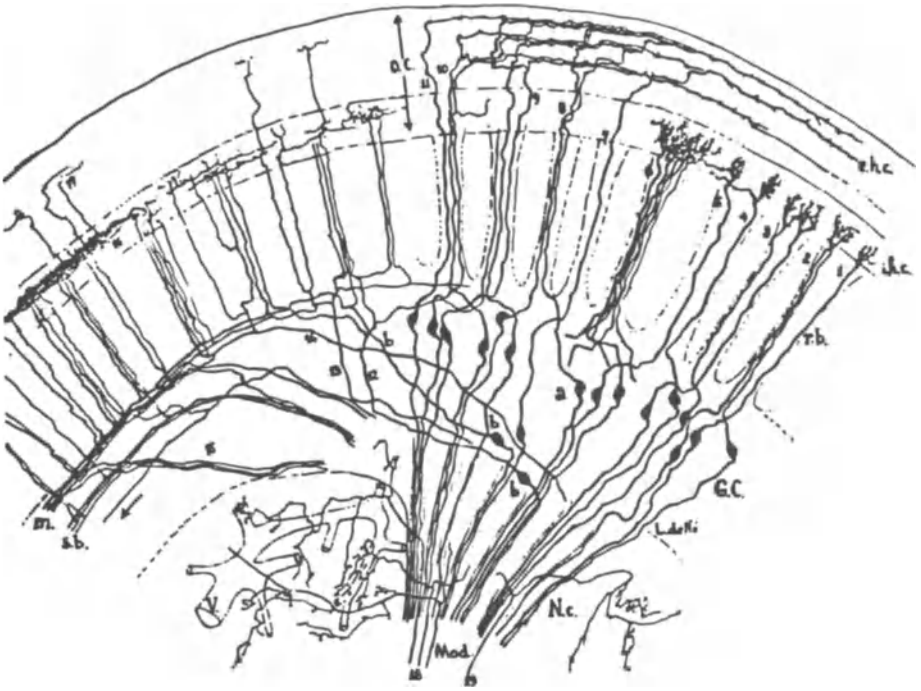


Figure 2. Nerve cells and fibers found in a single section that passes through the basal spiral turn of the cochlea (21).

and can be partly or even completely depolarized by sufficiently large stimuli. In particular, they respond to changes in the external potassium concentration by partial polarization or depolarization, with resultant changes of potential in the extracellular fluid comparable with those associated with an active neuron (i.e., up to several millivolts in amplitude), but on a time scale which is long compared with that of an action potential. By modeling the Debye layers associated with the glial membrane, we shall reproduce these observations and further show that they may serve at least one important purpose, other than that of regulating the concentration of potassium: helping to transmit electrical potentials through the extracellular medium.

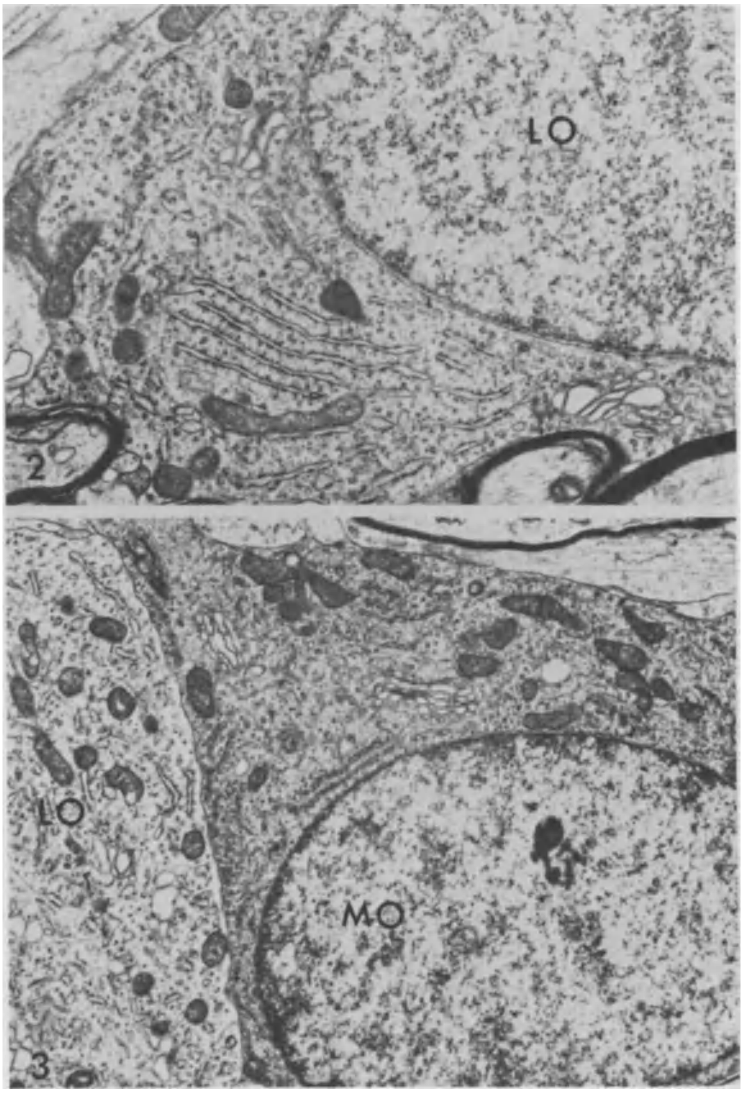


Figure 3. Oligodendrocytes in 7-week old rat (11).

In developing a model for extracellular activity we shall use a natural extension of the methods developed in previous papers (11-14) for the modeling of electrical activity in neurons. This work, based on the firmly established laws of electrostatics and ion dynamics, allowed the calculation of changes in electrical and chemical potentials in the Debye layers of the neural membrane during the course of graded and action potentials. It was shown that the temporal and spatial variations of these potentials were determined in principle by a dispersion equation which has resonances associated with ions present in low concentrations: calcium and sodium at the inner membrane surface. In this way the essential role of calcium in these processes could be explained, but contributions to action potentials from the extracellular layer, small compared with those at the inner membrane surface, were neglected. In the extracellular environment, however, these contributions (of the order of several millivolts) are of major importance. We shall show in the next section that they are determined to a large extent by variations in the concentration of calcium and potassium ions, which can be appropriately called the calcium and potassium resonances.

The role of calcium is similar to that identified for it at the inner membrane surface, but as we shall show, the period associated with the potassium resonance at its maximum is considerably longer than those of the calcium and sodium resonances. It is, in fact, of the same order as the observed period of the alpha-rhythm. But for reasons which will appear later, the relation between the frequency of the potassium resonance and the alpha-rhythm is not as direct at this might suggest; synchronization of potential changes at the outer membrane surfaces of both neurons and glial cells is required.

The function of this paper is to present our preliminary findings concerning the role of potassium ions in extracellular activity of various types. There is a very large amount of experimental evidence on this question, which has been admirably summarized in another review article by Somjen (17). Some of the more relevant facts, derived from many sources, are as follows. The concentration of potassium in the extracellular fluid of the cortex, and in the cerebrospinal fluid in general, is in a small range near 3 mM/l, significantly below that of the blood, and is very stable except under abnormal conditions such as anoxia. The diffusion constant of potassium measured in brain slices is only 1/20 that of ordinary aqueous solution, and considerably below that of intact brain. Such diffusion is far too slow to account for the transient changes associated with neuronal activity, so some unidentified form of active transport must occur. Deviations from Nernst's relation ($\phi - \phi_k$ = constant) across the membrane are rapidly reduced, except when the external potassium level is very much lower than normal. However, the response of glial cells to changes in external potassium concentration is not at all times in accordance with Nernst's relation, and is therefore not simply passive. Diffusion of potassium across glial membranes is abnormally rapid, except where glial sheaths intervene. These facts, many of which are puzzling at first sight, are immediatly intelligible in the light of the rapid transport of potassium associated with the potassium resonance in the extracellular environment.

The interest of this investigation is heightened by recurrent suggestions that potentials in the extracellular fluid may provide a means of communication and transfer of information between neurons additional to, and to a large extent independent of, the synaptic connections of the neural network. Marrazzi and Lorente de No (18) summarized and added to the experimental evidence already existing at that time, showing that (i) impulses traveling in a neural fiber may act as liminal stimuli and set up new impulses (graded or action potentials) in neighboring fibers; (ii) the passage of an impulse along a fiber or group of fibers may modify the electrical excitability of the neighboring fibers; and (iii) such interactions may serve to synchronize spontaneous firing rhythms and modify the speed of conduction in neural fibers. In the work already referred to, Rall (8), cited later experimental evidence and independently raised the question whether it was physically realistic to think in terms of a coupling between neurons through extracellular potential gradients. He concluded on the basis of a rough calculation that an extracellular potential gradient of 5 mV/mm could modify the somatic potential by as much as 1 mV, but pointed out that more detailed calculations were called for.

In the following sections, we shall develop a preliminary model of the extracellular field at the microscopic level. However, for some purposes it is desirable to have a corresponding macroscopic theory of the intra-cortical field. Because of the geometrical complexity of the extensive system of neurons and glial cells, such a theory must necessarily be of a statistical character. Since the system has short-range but no long-range order, a description similar to that of liquid structure (19,20) seems appropriate. In a subsequent paper we will present a statistical theory directed toward this end, a few preliminary ideas of which are outlined in the APPENDIX.

THE POTASSIUM RESONANCE

Here, as in our earlier work (13, 14), we shall make use of the following fundamental relations between the ionic number densities n_a, the chemical potentials ϕ_a, the ionic current densities $\underline{j}_a$, and the electrical potential ϕ :

$$
\begin{aligned}
n_a &= n_a^o \exp(-\beta e_a \phi_a) \\
\underline{j}_a &= e_a n_a \nabla(\phi_a - \phi)/\theta_a \\
dn_a/dt &= n_a \nabla\cdot(\underline{j}_a/e_a n_a) \\
\nabla^2\phi &= -(4\pi/\kappa)\Sigma_a e_a n_a
\end{aligned}
\tag{1}
$$

Where the subscript a takes the values K, Na, Ca and Cl, corresponding to the types of ions which we wish to consider. The n_a^o are values of the n_a at some undisturbed point, relative to which chemical potential differences are measured; $\beta = 1/(kT)$ is inversely proportional to the absolute temperature; the e_a and θ_a are ionic charges and resistances; and κ is the dielectric constant. Although the θ_a do vary somewhat with the degree of hydration of the ions and the polarization of the electrolytic fluid, they will be regarded as constants for phenomena on the time scale in which we are interested (10^{-3} sec to 10^{-1} sec). Thus the large variations in the conductances $e_a n_a/\theta_a$ postulated by Hodgkin and Huxley in their well-known model of the action potential, and by others following them, will be attributed to variations in the ionic densities in the Debye layers of the neural membrane.

From eq. 1, it is easy to derive the coupled set of nonlinear equations

$$
\gamma_a \frac{d\phi_a}{dt} = \nabla^2\phi_a + \Sigma_b C_b^o (e^{-\alpha_b \phi_b} - 1) \tag{2}
$$

Where the $\alpha_b = \beta e_b$ are just the ionic valencies if a suitable unit (24 mV) is adopted for the potentials, the $C_b^o = 4\pi\beta e_b n_b^o/\kappa$ are products of the valencies and the fixed concentrations (in mM if the unit of distance is 14.7 Å at a temperature of 310°K), and γ_a is the mean resistance of the membrane and its associated Debye layer for the a - th type of ions in the dynamical state (taken to be a constant). It follows from eq. 2 that an arbitrary small change $\delta\phi_a$ of the ϕ_a in time or position will satisfy the linear equation

$$
\gamma_a \frac{d(\delta\phi_a)}{dt} = \nabla^2(\delta\phi_a) - \Sigma_b \alpha_b C_b \delta\phi_b \tag{3}
$$

where

$$
C_b = C_b{}^o \exp(-\alpha_b \phi_b) = \pi \beta e_b n_b/\kappa
$$

If, as is often so, particularly in the extracellular fluid, the variations are independent of the ϕ_b, this equation can be solved by means of the ansatz

$$
\nabla^2(\delta\phi_a) = p \frac{d(\delta\phi_a)}{dt} + q\delta\phi_a \tag{4}
$$

Where P and q are, in general, functions of position. As in our earlier work, we are not interested in solutions which either tend rapidly to zero with time, or grow rapidly and therefore fail to satisfy the condition that $\delta\phi_a$ should be bounded. Consequently we look for periodic solutions and find, as before, that the admissible frequencies ω satisfy the dispersion equation

$$\Sigma_b \frac{\alpha_b C_b}{q + i\omega\,(p - \gamma_b)} = 1 \qquad (5)$$

since the denominator is complex, this resolves into two real equations connecting P, q and ω. These equations are solved most conveniently by substituting values of p within the admissible range (between γ_K and γ_{Cl}, where ω may be real), and determining the roots of the cubic in q^2/ω^2 which results from taking the imaginary part of eq. 5. The real part of this equation then determines the corresponding value of q^{-1} and hence of ω.

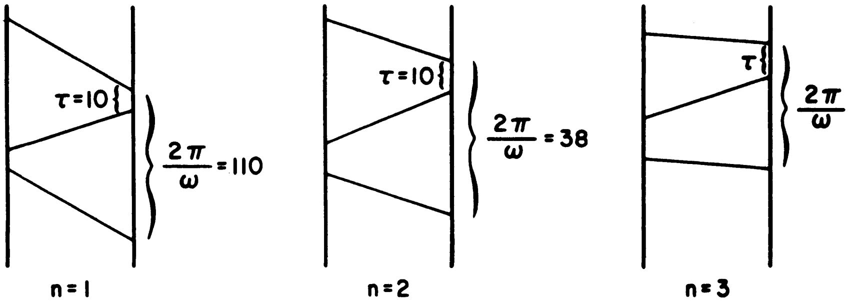


Figure 4. Possible eigenstates of potassium potential propagating in an extracellular neuron-glial cleft.

The general character of solutions for the inner membrane surface of certain neurons of interest in invertebrate neurophysiology has already been described in earlier papers (13, 14). For the neurons of the vertebrate cortex the solutions are quite similar and will not be discussed here, but it is worth noting that the concentration of calcium and sodium (C_{Ca} = .075 mM and C_{Na} = 12.5 mM) are considerably higher than in the work on invertebrates, and that the level of sodium is near the upper instead of the lower limit at which an action potential is possible. This suggests a regulating mechanism in which the elimination of sodium as a result of the metabolic processes of the cell is routinely offset by some form of activity tending to raise the sodium concentration. Action potentials are of course one such form of activity, but because of their occasional nature they must be supplanted by some other process during quiet periods. We are therefore led to consider the possibility of activity at the outer membrane surface. The concentrations in the extracellular fluid (C_K = 3 mM, C_{Na} = 150 mM, C_{Ca} = 1.25 mM, C_{Cl} = -125 mM) are very different from those considered previously, and the structure of the roots of eq. 5 is correspondingly different.

The resonances, as at the inner membrane surface, are associated with the ions at low concentrations, in this instance calcium and potassium ions, and are characterized by small values of q, which correspond to values of P just below γ_{Ca} and just above γ_K respectively. The properties of the potassium resonance, as will be shown, are not precisely analogous to those of the sodium resonance, partly because γ_k - P is negative, and partly because there is approximate electrochemical equilibrium between potassium (but not sodium) in the extracellular and intracellular fluids. Of special significance from our point of view are the remarkably small values of q and ω associated with the potassium resonance, which imply that both the thickness of the extracellular Debye layer and the time scale of potentials in this layer are much longer than those at the internal surface of the neural membrane. This is fully consistent with the experimental data referred to in the first section. In physical terms, the explanation is related to the fact that the degree of hydration of the potassium ions is less than that of other types of ions in the electrolyte, so that the resistance γ_K experienced by the potassium ions is also much smaller than the other γ_a.

The role of calcium in exciting the potassium resonance is similar to that described in relation to the sodium resonance at the inner membrane surface in our earlier work. A computer simulation much like that described before (14) shows that, because the calcium is in relatively greater abundance, the associated resonance is even more easily excited. The largest fluctuations are, of course, in the calcium concentration; but there are also fluctuations in the other ionic concentrations. In the case of potassium these are sufficient to excite resonant states with amplitudes in the nonlinear domain.

The potassium resonance may also be excited by the external fluctuations in the sodium concentrations which follow an action potential in a neuron, or by fluctuations in electrical potential in the extracellular fluid. Moreover, small fluctuations in the potentials at the surface of a glial cell will excite this resonance in exactly the same way as at the surface of a neuron, establishing that such activity is by no means limited to the neural network.

EXTRACELLULAR ELECTRICAL ACTIVITY

We wish to consider the propagation of a potential between an axon and a glial cell, between two glial cells, and between a glial cell and the dendrite of a neuron. It is expected that this activity will be associated mainly with the potassium resonance, since this is the only resonance with an amplitude of more than one mV in the extracellular region. The sodium resonance will not extend far outside the membrane, while the calcium resonance will be of small amplitude. Denoting by ψ_a the contribution to the potentials ϕ_a from the potassium resonance, the equation

$$\nabla^2\psi_a = p\,\frac{\partial\psi_a}{\partial t} + q\psi_a \tag{6}$$

must be satisfied, where $p = \gamma_K \simeq 1.08$ and $q^{-\frac{1}{2}}$, the thickness of the associated Debye layer is of the order of 1000 Å.

If z and r are the longitudinal and radial coordinates,

then
$$\frac{\partial\psi_a}{\partial z} = -\frac{1}{v_z}\frac{\partial\psi_a}{\partial t} - k_z\,\psi_a$$

and
$$\frac{\partial\psi_a}{\partial r} = -\frac{1}{v_r}\frac{\partial\psi_a}{\partial t} - k_r\,\psi_a \tag{7}$$

where v_z and v_r are the longitudinal and radial velocities of propagation, and k_z and k_r are the longitudinal and radial attenuation coefficients. The signs adopted assume propagation in the increasing z- and r-directions; all signs are reversed on the right hand side for propagation in the decreasing r-direction. Assuming v_z to be much larger than v_r and k_z to be substantially smaller than k_r, as it must be if the potential is to propagate 1 mm or more between nodes along the axon, substitution in equ. 6 reduces to

$$\frac{1}{v_r^2}\frac{\partial^2\psi_a}{\partial t} + \frac{2k_r}{v_r}\frac{\partial\psi_a}{\partial t} + k_r^2\,\psi_a = p\,\frac{\partial\psi_a}{\partial t} + q\,\psi_a \tag{8}$$

At small amplitudes, and also to a good approximation in the non-linear region, the potential is time reversible; thus we have

$$-\,\omega^2/v_r^2 + k_r^2 = q$$

$$2k_r/v_r = p \tag{9}$$

where ω is the frequency deduced from the linearized theory. As p, q and ω are known, the above equations determine v_r and k_r :

$$v_r = 2k_r/p$$

$$k_r^2 - p^2\omega^2/4k_r^2 = q \tag{10}$$

Obviously $k_r > q^{\frac{1}{2}}$, and if q were not very small the radial propagation would be attenuated so strongly that it could not reach between an axon and a glial cell.

If $\psi_a = f_a(t)$ for $r = r_o$ on the outer surface of the neuron (or glial cell), the solution of equ. 8 is

$$\psi_a = e^{-k_r(r - r_o)} f_a \left(t - \frac{r - r_o}{v}\right) \tag{11}$$

for $r > r_o$ in the extracellular fluid. To this we should add a solution propagating in the opposite direction, corresponding to the potential evoked at the external surface of the opposite cell:

$$\delta\psi_a = e^{-k(r_o' - r)} f_a' \left(t - \frac{r_o' - r}{v}\right) \tag{12}$$

where $r_o' - r = d$ is the separation of the two external surfaces and $f_a'(t)$ is the evoked potential at $r = r_o'$.

For small amplitudes

$$f_a(t) = f_a(o) \cos(\omega t)$$

$$f_a'(t) = f_a'(o) \cos[\omega(t - d/v_r - \delta)] \tag{13}$$

where d/v_r represents the time of propagation of the potential between the two surfaces, and δ is the delay in response of the glial cell, which will depend on the amplitude of the stimulus and must be determined from the computer program mentioned earlier.

For $r + (r_o = r_o')/2$, the most likely position of a microelectrode halfway between the two surfaces, this gives

$$\psi_a(t) = e^{-\frac{1}{2}k_r d} \left\{ [f_a^o \cos[\omega(t - d/2v_r - \delta/2)] + [f_a'' \sin[\omega(t - d/2v_r - \delta/2)] \right\} \tag{14}$$

where f_a^o and f_a'' depend on the relative magnitudes of f_a and f_o'.

Cylindrical or spherical symmetry would no doubt be more appropriate than the plane symmetry assumed in the above analysis, but this could not change the results by more than a factor of order $k_r d$. In any event, it is clear that potential variations of the type observed with invasive electrodes and presented in Figure 1 could result from the superposition of terms in equ. 14.

SOLUTIONS IN THE EXTRACELLULAR REGION

Our method of solving

$$\gamma_a \frac{\partial \phi_a}{\partial t} = \frac{\partial^2 \phi_a}{\partial x^2} + \sum_b C_b^o \left(e^{-\alpha_b \phi_b} - 1 \right) \tag{15}$$

in the extracellular region, where the C_b^o are constants, is essentially as follows: we write

$$\phi_a = \int_{x_\ell}^{x} X_a^o \, dx + \int_{x}^{x_\ell + d} X_a^g \, dx \tag{16}$$

where X_a^o is the contribution to the potential from the neuron (or glial cell) and X_a^g the contribution from its neighbor at distance d (200 Å). Then we may also write

$$\frac{\partial \phi_a}{\partial x} = X_a^o + X_a^g = X_a \tag{17}$$

By differentiating equ. 15 with respect to x we get

$$\gamma_a \frac{\partial X_a}{\partial t} = \frac{\partial^2 X_a}{\partial x^2} - \sum_b \alpha_b \, C_b \, X_b$$

$$C_b = C_b^o \, e^{-\alpha_b \phi_b} \tag{18}$$

Note that $X_a(x,t)$ depends on x only through the ambient 'concentrations' (actually ambient charge densities) C_b, which in turn depend on values of X_b^o (x′, t) for x′ < x and values of X_b^g (x′, t) for x′> x. We can regard the C_b as already determined; our task is to determine X_a. We begin by writing

$$X_a(x,t) = \int_{-\infty}^{\infty} \overline{X}_a(p,x,t) \, f(p) \, dp \tag{19}$$

where

$$\frac{\partial^2 \overline{X}_a}{\partial x^2} = p \frac{\partial \overline{X}_a}{\partial t} + q \, \overline{X}_a \, , \; q + q(p) \tag{20}$$

using the ansatz. From equ. 18 we then have

$$(\gamma_a - p) \frac{\partial \overline{X}_a}{\partial t} = q \, \overline{X}_a - \sum_b \alpha_b \, C_b \, \overline{X}_b \tag{21}$$

from which it follows in the usual way that

$$\overline{X}_a = R \frac{fe^{i\omega t}}{(p-\gamma_a)i\omega+q} \tag{22}$$

where f is a complex 'constant' of integration (which actually depends on x) and q and ω are determined in terms of p and the concentrations C_b by

$$\sum_b \frac{\alpha_b C_b}{(p-\gamma_b)i\omega+q} = 1 \tag{23}$$

-- a complex equation equivalent to two real equations. To satisfy boundary conditions and also to insure stability, p must be restricted to ranges between γ_K and γ_{Na}, and between γ_{Na} and γ_{Ca}, in the usual way.

To determine the x-dependence of f in reasonable approximation we neglect the relatively slow variation of the concentration C_b. with x, so that according to equ. 20 we must have

$$\frac{d^2f}{dx^2} = (pi\omega+q)f \tag{24}$$

and

$$f = f^o + f^g$$

$$f^o = k^o e^{-\mu(x-x_\ell)} , \quad f^g = k^g e^{\mu(x-x_\ell - 2d)}$$

$$\mu^2 = (pi\omega+q) \tag{25}$$

together with the condition that ϕ should vanish at the glial membrane.

The result of the above analysis is to show that, to a reasonable approximation (treating the C_b as constants neglects certain minor nonlinearities) potentials are attenuated as $\exp(-\mu x)$, independent of amplitude, in the extracellular region.

We can therefore suppose that the potential of the glial cell membrane surface satisfies

$$(\gamma_a - p)\frac{d\phi_a^g}{dt} = q\,\phi_a^g + \sum_b C_b^o\left(e^{-\alpha_b(\phi_b^g + \lambda\,\phi_b^o)} - 1\right) \tag{26}$$

while that of the neural membrane surface satisfies

$$(\gamma_a - \bar{p})\frac{d\phi_a^o}{dt} = \bar{q}\,\phi_a^o + \sum_b C_b^o\left(e^{-\alpha_b(\phi_b^o + \lambda\phi_g^o)} - 1\right) \tag{27}$$

where $\lambda = e^{-|\mu|d}$ represents the attenuation factor, ignoring for the moment any indirect activity of the neuron.

In first order nonlinear approximation, these last equations yield the relations

$$(\gamma_a - p)\frac{d\phi_a^g}{dt} = q\phi_a{}^g - \sum_b \alpha_b(\phi_a{}^g + \lambda\phi_b{}^o)[1 - \frac{1}{2}\alpha_b(\phi_a{}^g + \lambda\phi_b{}^o)]$$

$$(\gamma_a - \bar{p})\frac{d\phi_a{}^o}{dt} = \bar{q}\phi_a - \sum_b \alpha_b (\phi_a{}^o + \lambda\phi_b{}^g)[1 - \frac{1}{2}\alpha_b(\phi_b{}^o + \lambda\phi_b{}^g)] \qquad (28)$$

-- coupled nonlinear equations of the van der Pol type. We expect these equations to have periodic solutions with a definite amplitude, as well as attenuated solutions and growing solutions which must be rejected as not satisfying boundary conditions. We also expect to find a solution with a frequency near ω and another with a frequency near 2ω, reflecting the presence of α- and β-waves.

The value of λ in equ. 28 depends quite sensitively on the real part of μ , which in turn depends sensitively on ω . Also, the response time τ of the glial cell to a signal from the neuron depends on the value of λ. Thus τ is a function of ω, and it may be argued that $\tau(\omega)$ must be a multiple of π/ω , i.e., that there exists an eigenvalue equation of the form

$$\tau(\omega) = \frac{n\pi}{\omega} \qquad (29)$$

for determining ω.

The argument is that, to obtain stable oscillations, the time required for the glial cell to respond must be half, or some multiple of one-half, the period of the oscillation. The time of propagation of a signal between a glial cell and a neuron could be negligible compared with the response time (even at a very low velocity of propagation, 200 Å would be covered in a small fraction of a μsec); and on this assumption, it is necessary for the response time to be in phase with the oscillations to avoid constructive (leading to instability) or destructive (leading to rapid quenching) interference.

Generalizing the above argument, the eigenvalue condition becomes a dispersion equation, which permits a small range of frequencies for both α- and β-waves, and also allows for the propagation of these waves in the longitudinal as well as the transverse direction. The time of propagation between the adjacent surfaces then becomes significant. Instead of equ. 15 we must write

$$\gamma_a \frac{\partial\phi_a}{\partial t} = \frac{\partial^2\phi_a}{\partial x^2} + \frac{\partial^2\phi_a}{\partial z^2} + \sum_b C_b^0 \left(e^{-\alpha_b\phi_b} - 1\right) \tag{30}$$

since the velocity of propagation in the longitudinal direction may be considerably less in the α- and β-modes than for an action or a graded potential. If v_z is the velocity of propagation in the z-direction,

$$\frac{\partial^2\phi}{\partial z^2} = \frac{1}{v_z^2} \frac{\partial^2\phi_a}{\partial t^2} \tag{31}$$

If we retain the ansatz (equ. 20), equation 21 becomes

$$(\gamma_a - p) \frac{\partial\overline{X}_a}{\partial t} = \frac{1}{v_z^2} \frac{\partial^2\overline{X}_a}{\partial t^2} + q\, \overline{X}_a - \sum_b \alpha_b C_b \overline{X}_b \tag{32}$$

Then equ. 22 is changed to

$$\overline{X}_a = R \frac{fe^{i\omega t}}{(p-\gamma_a)i\omega+\overline{q}} \tag{33}$$

where

$$\overline{q} = q - \frac{\omega^2}{v_z^2} \tag{34}$$

and q satisfies the same equation as q,

$$\sum_b \frac{\alpha_b C_b}{(p-\gamma_a)i\omega+\overline{q}} = 1\ , \tag{35}$$

when ω^2/v_z^2 is neglected. Obviously the effect of a small velocity of propagation v_z is to increase the value of q. This will not significantly reduce the thickness of the Debye layer,

$$a_D = \mu_r^{-1} = [\tfrac{1}{2} q + \tfrac{1}{2}(\omega^2p^2 + q^2)^{\frac{1}{2}}]^{-\frac{1}{2}}\ , \tag{36}$$

provided

$$v_z^2 \gtrsim \frac{\omega}{p} \tag{37}$$

The radial velocity of propagation is

$$v_x = 2\mu_r/p = [2q + 2(\omega^2p^2+q^2)^{1/2}]^{1/2}/p \tag{38}$$

which depends only weakly on q (because $\omega^2p^2>q^2$), and hence on the wave vector $k = \omega/v_z$ of the longitudinal motion, so long as the condition in equ. 37 is satisfied. From this it appears that $v_x = (2\omega/p)^{\frac{1}{2}}$ is a good approximation. The time of transit across the external cleft between the neuron and glial cell must therefore be $(p/2\omega)^{1/2}d$, and for $d \simeq 15$ (or 150 Å) this proves to be of the same order as the response time. Consequently, both must be taken into account in formulating the desired dispersion equation.

Adding the time of propagation $(p/2\omega)^{1/2}d$ to the response time $\tau(\omega)$ and, by the same reasoning as before, equating this to a multiple of the half period, we obtain the generalized eigenvalue equation

$$(p/2\omega)^{1/2}d + \tau = n\pi/\omega \tag{39}$$

or more accurately, the dispersion equation

$$d/v_x + \tau = n\pi/\omega \tag{40}$$

with v_x given by equ. 38. This latter form shows the dependence on the wave vector $k = \omega/v_z$ contained in $q = \bar{q} + k^2$, but the amount of dispersion is clearly quite small.

Within the 40-100 msec range, τ is rather insensitive to the period $T = 2\pi/\omega$, and has a value of about 10 msec. If $n = 1$ we obtain from equ. 39, $T \simeq 20 + 15(T/\pi)^{1/2} \simeq 110$ msec, but if $n = 2$ the result is $T = 10 = 15\ (T/4\pi)^{1/2} \simeq 38$ msec (see Figure 4). These values are in reasonable agreement with reported values for the α- and β-waves. For $n = 3$, however, the time of transit is of the same order as the response time, and it is difficult to discern a definite period.

It should be noted that the boundary conditions of a cell which is active (permeable to potassium) are different from those of an inactive membrane containing electrolytic fluid. For the latter, both ϕ and $\frac{\partial\phi}{\partial x}$ would have the same value a the two surfaces. The glial cell will also have the same value of $\frac{\partial\phi}{\partial x}$ at its two surfaces, but will acquire a surface electric dipole moment which varies in a rather complex way with ϕ^o, the field propagating from an extended source. The effect of this dipole moment is to produce a potential difference across the membrane surface equal to $\phi\delta$, as determined by our nonlinear equation. In the nonlinear domain, $\phi^o + \phi\delta = 0$, when $\phi\delta$ attains its maximum, just outside the membrane surface. In the linear domain, resonant behavior produces slightly nonlinear oscillations.

Additional to these macroscopic boundary conditions, there are, of course, microscopic conditions which must be applied to obtain stable solutions of the nonlinear equations. These and the foregoing boundary conditions will be described in more detail in a later paper.

REFERENCES

1. P. Anderson and S.A. Andersson, Physiological Basis of the Alpha Rhythm, Appleton-Century-Crofts, New York (1968).
2. W.J. Freeman, Mass Action in the Nervous System, Academic Press, New York (1975).
3. F.H. Lopez da Silva and W. Storm van Leeuwen, Neurosci. Lect. 6:237-241 (1977).
4. E. Bas̆ar, EEG-Brain Dynamics, Elsevier, Amsterdam (1980).
5. R. Lorente de No, A Study of Nerve Physiology II: Studies from the Rockefeller Institute for Medical Research 132:1-543, Rockefeller Inst. for Med. Res., New York (1947).
6. R. Plonsey, Bioelectric Phenomena, McGraw-Hill, New York (1968).
7. R. Plonsey, in: Biomagnetism, S.N. Erne, H.D. Hahlbohm and H. Lubbig (eds.), DeGruyter, Berlin, pp. 177-205 (1981).
8. W. Rall, Biophys. J. 2:145-167 (1962)
9. A.L. Towe, in: Bioelectric Recording Techniques A, R.F. Thompson and M.M. Patterson (eds.) Academic Press, New York (1973).
10. R. Lorente de No, The Primary Acoustic Nuclei, Raven Press, New York (1981).
11. J. A. Paterson, in: Eleventh International Congress of Anatomy, A, E.A. Vidrio and S. Federoff (eds.) Alan R. Liss, Inc., New York (1981).
12. G. Somjen, Ann. Rev. Physiol. 37:163-190 (1975).
13. T. Triffet and H.S. Green, J. Theor. Biol. 86:3-44 (1980).
14. T. Triffet and H.S. Green, J. Theor. Biol. 100:645-675 (1983).
15. H.S. Green and T. Triffet, Math Modeling, 1:41-61 (1980).
16. H.S. Green and T. Triffet, Math Modeling, 3:161-178 (1982).
17. G.G. Somjen, Ann. Rev. Physiol. 41: 159-177 (1977).
18. A.S. Marrazzi and R. Lorente de No, J. Neurophysiol. 7:83-102 (1944).
19. H.S. Green in: Handbuch der Physik X, S. Flugge (ed.), Springer, Berlin, pp. 1-133 (1960).
20. H.S. Green, Molecular Theory of Fluids, Dover, New York (1969).
21. D. W. Cayley, in: Cellular Aspects of Neural Growth and Differentiation, D.C. Pease (ed.), U. of California Press, Berkeley pp. 73-102 (1971).

APPENDIX

Since the extracellular environment has no long-range structure, a statistical description somewhat analogous to that used for liquid structure seems appropriate. The mathematical formulation is in terms of a set of probabilities $\nu_r(\underline{X})$, $\nu_{rs}(\underline{X},\underline{X}')$, $\nu_{rst}(\underline{X},\underline{X}',\underline{X}'')$, ... where r, s, t, ... take 3 values: e for extracellular fluid, n for neurons and g for glial cells. Then $\nu_e(\underline{X})$

denotes the probability that the point $\underline{X}$ is in the extracellular fluid, $\nu_n(\underline{X})$ the probability that it is in a neuron and $\nu_g(\underline{X})$ that it is in a glial cell while $\nu_{eg}(\underline{X},\underline{X}')$ denotes the probability that $\underline{X}$ is in the fluid and $\underline{X}'$ in a glial cell, etc. Obviously

$$\sum_r \nu_r(\underline{X}) = 1$$

$$\sum_s \nu_{rs}(\underline{X},\underline{X}') = \nu_r(\underline{X}),$$

and

$$\sum_t \nu_{rst}(\underline{X},\underline{X}',\underline{X}'') = \nu_{rs}(\underline{X},\underline{X}')$$

and so on. Also, because of the absence of long-range order $\nu_{rs}(X,X') = \nu_r(\underline{X})\nu_s(\underline{X}')$ for large $|\underline{X} - \underline{X}'|$, $\nu_{rst}(X,X',X'') = \nu_r(\underline{X})\nu st(\underline{X}',\underline{X}'')$ for large $|\underline{X} - \underline{X}'|$ and $|\underline{X} - \underline{X}''|$, ... Obviously, $\nu_e(\underline{X})$ is just the fraction of extracellular volume, $\nu_n(\underline{X})$ is the fraction of neural volume, ...; and these can be determined from observations. For small values of $|\underline{X} - \underline{X}'|$, $|\underline{X}' - \underline{X}''|$, ... the functions $\nu_r(\underline{X},\underline{X}')$ and $\nu_{rs}(\underline{X},\underline{X}',\underline{X}'')$ determine the short-range order, or structure of the system.

Other statistical properties can be determined from the ν-functions. For instance, since $\nu_{en}(\underline{X},\underline{X})$ obviously vanishes, for small $\underline{\delta}$

$$\nu_{en}(\underline{X}, \underline{X} + \underline{\delta}) = \nu_{en}(\underline{X}, \underline{X} + \underline{\delta}) - \nu_{en}(\underline{X},\underline{X})$$

$$= \underline{\delta} \cdot \left| \frac{\partial}{\partial X'} \nu_{en}(\underline{X},\underline{X}') \right|_{\underline{X}' = \underline{X}}$$

Thus if

$$\underline{P}_{en}(\underline{X}) = \frac{1}{\nu_e(\underline{X})\nu_n(\underline{X})} \left| \frac{\partial}{\partial \underline{X}'} \nu_{en}(X,X') \right|_{\underline{X}=\underline{X}'}$$

$\underline{\delta}\cdot\underline{P}_{en}(\underline{X})\ {}_n(\underline{X})$ is the probability that a small displacement δ will move from the point $\underline{X}$ in the extracellular fluid across the membrane of a neuron. Then the direction of $\underline{P}_{en}(\underline{X})$ defines the normal to the neural membrane, while its magnitude is the increase of the distance of separation of the neuron and the neighboring (glial) cell. These functions can also be assumed known form observation; they determine the geometrical characteristics of the intracellular environment.

The ν-functions are obviously analogs of the number distribution functions n_a, n_{ab}, n_{abc}, ... of the ionic theory of electrolytes transferred to this purely geometrical context. To determine the extracellular field on analog of the velocity distribution functions f_a, f_{ab}, f_{abc}, ... is needed. Calling these $_{ra}$ $_{ra,sb}$ ...., where $\psi_{ra} = \psi_{ra}(\underline{X}, \phi_b, t)$ satisfies a linear differential equation similar to that satisfied by X_a above, then

$$\gamma_a \frac{\partial \Psi_{ra}}{\partial t} = \nabla^2 \Psi_{ra} - \sum_b \alpha_b C_b^o e^{-\alpha_b \phi_b} \Psi_{rb}$$

determines the macroscopic potentials $\bar{\phi}_a(\underline{X}, t)$ and $\tilde{\phi}_a(\underline{X}, t)$ through the equation

$$\nu_e(\underline{X}) \bar{\Phi}_a (\underline{X},t) = \alpha_a \int \Psi_{ea}(\underline{X}, \phi_b, t) \, d^2\phi$$

for the extracellular fluid, or the equation

$$\tilde{\phi}_a(\underline{X}, t) = \sum_r \alpha_a \int \Psi_{ra} (\underline{X}, \phi_b \; t) \, d^2\phi$$

for the inhomogeneous medium as a whole.

FREQUENCY SELECTIVE PROPAGATION OF EXTRACELLULAR ELECTRICAL STIMULI TO INTRACELLULAR COMPARTMENTS

H.P. Schwan

Department of Bioengineering
D-3 University of Pennsylvania
Philadelphia, Pennsylvania 19104

INTRODUCTION

Considerable efforts have been made during the past two decades to explore the relationship between ambient field strength and flux values of electromagnetic fields and field strength values or specific absorption rates (SAR) in man and animals (1). These macroscopic studies extend over the total frequency range from very low frequencies to the microwave range and established the "coupling" mechanism which relates *in situ* to external ambient values.

In contrast, virtually nothing has been done to relate local *in situ* field values to field values in the various cellular compartments. These include the outer membrane and its extended "fuzzy" surface structures, membranes of subcellular organelles, such as mitochondria and cell nucleus, membrane structures connecting with the outer membrane such as the tubular system and endoplasmic reticulum and the various interior compartments surrounding proteins, nucleic acids and the cytoskeleton. This more microscopic task becomes now increasingly urgent as efforts increase to formulate a possible nonthermal field interaction mechanism at the cellular and macromolecular level. Clearly, in order to quantitatively apply whatever direct interaction mechanism that may be of interest to any particular cellular compartment, one must utilize the relationships which connect *in situ* field values or field values in the extracellular medium to values in the compartment under consideration. We will present an outline of some relationships already established, and indicate others requiring further work. All sorts of strong frequency dependencies will be noted, quite distinct for each compartment. We speculate, therefore, that any nonthermal effect should be highly frequency dependent for a given internal field or SAR value throughout the total RF and ELF range. Finally, we will briefly summarize some nonthermal effects which have been of rapidly increasing interest in recent years.

Membrane Potential

For a spherical cell of radius R in a medium with the field strength E, the induced membrane potential is (2):

$$Vm = \frac{1.5ER}{\sqrt{1 + (\omega T)^2}} \tag{1}$$

ω angular frequency and with the time constant T

$$T = RC(\rho_i + 0.5\,\rho_a) \tag{2}$$

where C the membrane capacitance per membrane unit area near 1 $\mu F/cm^2$ and the specific resistances of intracellular and extracellular media are ρ_i and ρ_a of the order of 10 mS/cm (3). The effect of the membrane conductance can be neglected (3). The cutoff frequency $fc = \frac{1}{2}\pi T$ for typical cell sizes is therefore in the upper KHz to lower MHz range. For E ~ 1 V/cm the low frequency membrane potential is in the mV range for typical cell sizes and proportional to radius R. The high frequency potential $V\infty$ is independent of R.

$$V_o = 1.5ER; \; V\infty = 1.5E/\omega C(\rho_i + 0.5\,\rho_a) \tag{3}$$

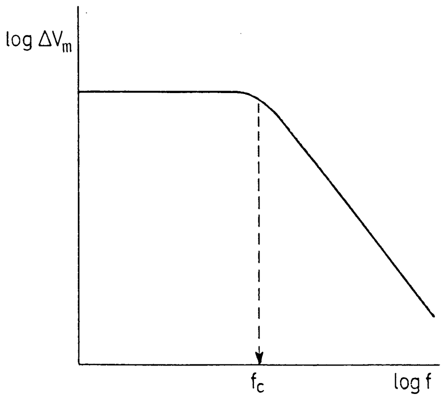


Figure 1. Frequency dependence of induced membrane potential.

Equations 1 and 2 and the conclusions derived from them, tacitly assume frequency independence of the membrane capacitance C. However, the apparent membrane capacitance changes at low frequencies for a variety of reasons, associated with connecting membrane systems and membrane environment. In addition, the membrane itself is likely to display capacitative and related conductive changes (5). Experimentally observed dispersive changes require the introduction of inductive elements in equivalent circuits. These inductive elements in combination with the capacitative components can give rise to resonances, so that the admittance magnitude can become large for certain combinations of frequency and applied DC potential. These singularities may well affect our conclusion of a frequency independent membrane potential below cutoff frequency fc. Almost nothing has been done to investigate this problem so far.

Intracellular Field Strength

The intracellular field strength and potential must sum vectorially with the membrane potential to account for the total potential drop across the cell, which is 1.5 ER. The intracellular field strength is given by:

$$E_i = 1.5ER \frac{\omega T}{\sqrt{1 + (\omega T)}} \qquad (4)$$

if intracellular and extracellular resistivities are equal. If they are not, a somewhat more complicated expression results. The frequency dependence is obviously that of a low pass filter with the cutoff frequency identical to that of the membrane potential.

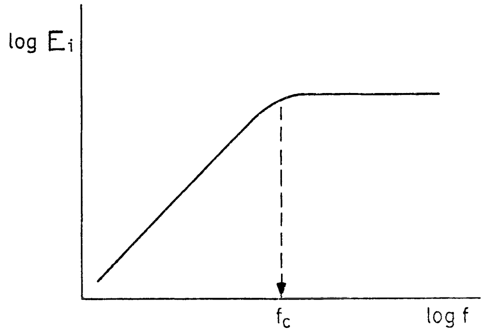


Figure 2. Frequency dependence of intracellular field strength for a constant extracellar field strength.

Well below this cutoff frequency, the intracellular space is well shielded, and the external field is not expected to interact with internal macromolecular content and organelles. Above the cutoff frequency internal field strength rapidly approaches values comparable to those in the extracellular space. Thus the radiofrequency range is characterized by a shift from more likely membrane interactions at the lower frequencies to interactions with components of the intracellular medium.

Membrane Potential of Nucleus and Organelles

If the internal field strength would be frequency independent, subcellular membrane potentials would obey the same rules outlined for the outer membrane. However, the internal field decreases with decreasing frequency as stated above. Thus, the membrane potential of organelles is obtained by multiplying the expressions 1 and 4

$$V_o = 2.25RE \frac{\omega T_2}{\sqrt{(1+\omega T_1)(1+\omega T_2)}} \qquad (5)$$

where T_1 and T_2 are time constants as given by equation 2, with radii equal to that of the cell and the organelle. The membrane potential is fairly frequency independent between two frequencies whose ratio is given by the ratio of linear cell and organelle dimensions. It drops off outside this range and is identical with the outer membrane potential at high frequencies.

Field Strength Inside Organelles and Nucleus

The medium inside organelles is double shielded from the external field by cellular membrane and organelle membrane. Hence, it is exceedingly weak at low frequencies, slowly gains above the cutoff frequency of the cells and increases rapidly above the cutoff frequency of the organelle, and shortly thereafter approaches values comparable to those in the extra- and intracellular space.

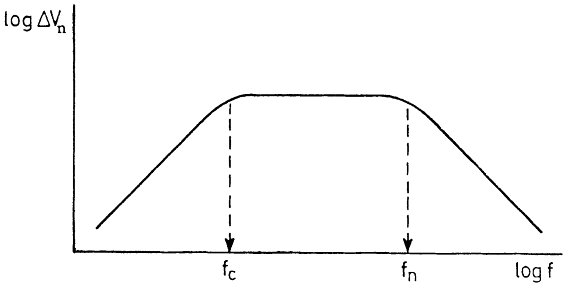


Figure 3. Frequency dependence of the induced nuclear membrane potential for a constant extracellular field strength.

Intracellular Membranes Connecting with the Outer Membrane

The potential across such membrane structures may be established by modeling these systems following a suggestion by Falk and Fatt (4). The inner membrane system contains a medium which is in contact with the outer cellular medium through a small "access" impedance of primarily resistive character. It thus appears as a large capacitor in series with the access impedance, the total series combination in parallel with the outer membrane. The cutoff frequency of this system is estimated to occur at frequencies below 1 KHz and thought to account in good part for the low frequency - dispersion of the dielectric properties of tissues (3, 5). Hence, the intracellular membrane systems considered here experience membrane potentials identical with those of the outer membrane only at very low frequencies and rapidly decreasing potentials at frequencies above the α-dispersion range.

The space enclosed by these connecting membrane systems is exposed only at very low frequencies to field values similar to extracellular values. As the frequency increases above the α -dispersion range the "access" impedance consumes most of the available potential and the membrane surrounded space appears largely field free. On the other hand the access to the membrane surrounded space is subject to nearly the same high field values imposed on the outer membrane.

An extension of these considerations is necessary to deal with the complexity of interconnecting membrane systems. Only a few epidermal systems, such as the gall bladder epithelium, have been investigated to a sufficient extent to propose equivalent networks. From the data provided by Schifferdecker and Froemter (6) it appears that frequency dependencies for the field in various components of these highly complex structures extend through the low frequency range up to about 20 KHz. Components include various cellular membrane systems, gap junctions and intracellular media. No attempt has been made yet to analyze fully the potentials and field strength values in all of these components as functions of frequency.

Outer Membrane Region

We shall neglect the possible role of membrane-bound water. Protein-bound water extends to a distance of only a few angstra. It, therefore, should contribute little to the effective dielectric properties of cells. It also appears not to hinder effectively ionic transport through the membrane. These arguments are fairly qualitative and rely on the similarity of membrane bound and established protein bound water.

It is well known that membranes, about 100 Å thick, are surrounded by brush-like glycoprotein structures which extend another 100 Å or more. Fixed charge densities of about one elementary charge per 10 square Å have been repeatedly quoted in the literature. Does this extended fixed charge distribution and associated counter ion cloud affect the potential distribution about the cell?

Schwarz has calculated the amplitude and relaxation frequency of the low frequency dielectric dispersion resulting from the movement of the counter ions with the alternating field (7). Using these expressions one obtains readily only a small contribution to the effective dielectric properties of typical tissue cells. This would imply that the outer cell structures do not materially affect the fields reaching the cell itself. However, it has been established that the ionic concentration profile near the cell surface and its fixed charges consumes a significant fraction of the potential applied to the total membrane complex (8). This would imply that the cell proper is significantly shielded by the fixed charge influenced ionic environment. Virtually nothing is known about the likely frequency dependence of this effect. If the pronounced low frequency-dispersion effects reported for a variety of cells are related to counter ion movement instead of cellular membrane invaginations, the shielding effect should vary accordingly through the low frequency dispersion range. The mechanism responsible for the cellular-dispersion data have not yet been entirely clarified (5) and the applicability of the Schwarz model is questionable, since it assumes a homogeneous distribution of fixed charges at the very surface of the cell. The effect of the distributed counter ion atmosphere near membranes on the potential distribution of cellular compartments remains, therefore, elusive.

Macromolecules

The influence of membrane bound water is estimated to be small as indicated above. However, it has a marked effect on the field which reaches proteins. This is caused by the relatively large amount of hydration of such macromolecules, typically in the range of about 0.3 g bound water per gram protein. Protein bound water is believed to undergo a marked relaxation, centered above some hundred MHz and characterized by a fairly broad distribution of relaxation times (9, 10). Through this frequency range the dielectric constant of this bound water decreases from near 80 to a few units. Therefore, the macromolecule is partially shielded from the external field in a frequency dependent fashion and depends on all the parameters involved.

In the radio frequency range proteins display another relaxation effect, most likely to be caused by their permanent dipole moment (11). The total hydrated molecule attempts to oscillate with the alternating field. No particular shielding effects need be considered since the successful theoretical treatment considers the rotation of the hydrated protein, i.e. the bound water is believed to move with the protein.

The dielectric properties of nucleic acids are more complex and only in part understood (12). It is therefore, not possible to model their frequency response and to apply the comments made about proteins.

Analytical Tools

The mathematical tools needed to determine field strength distributions at the cellular level have existed for some time and have been successfully applied to analyze the electrical properties of cell suspensions and tissues. They are essentially based on two theorems first presented by Maxwell (13). The first one states the conductivity of a suspension of particles of a given conductivity in a medium of different conductivity. The second states that it is possible to replace a particle surrounded by a shell by an "equivalent homogeneous" particle so that the external field remains the same. It states the equivalent conductivity as a function of shell and particle core conductivities. These relationships can be readily extended to the complex case by replacing real conductivities by complex ones. The exact derivation of the second problem is based on the solution of the Laplace potential equation for a spherical particle with a shell and provides, at the same time, the fields in each compartment.

The process of replacing a shell surrounded particle by an equivalent homogeneous one can be repeated in reducing an n-shell particle step by step to one. It has been shown that a suspension of shell surrounded particles can be stated precisely as a sum of two Debye-type relaxation expressions (14). By extension of this result the electrical behavior of an n-shell surrounded particle suspension can be written as a sum of n + 1 relaxations. All this can be cast in terms of equivalent networks. A formulation of the multiple shell case in terms of a finite ladder network was first given by Fricke (15). Simple closed form expressions for the dispersion terms and equivalent network elements for the one shell case have been provided by Schwan, (3) Schwan and Morowitz, (16) and Pauly and Schwan (14). An improved mixture theory for higher cellular volume concentration has been given by Hanai (17). A number of numerical approaches utilizing computers have been developed. Irimajiri, for example, used such a technique to analyze cells with nuclei (18).

The dependence of the induced membrane potential on the external field as stated in Equation 1 has been presented several times before with the conclusion that microwave field induced membrane potentials are for typical exposure conditions only in the range or less (19, 20). More recently, Drago et al., have undertaken a rigorous treatment investigating a number of cellular field strength values of interest, using a multi-component model of the cell including eight shells (21, 22). These papers are much more detailed than this one in stating the mathematical approach, obviating the need for doing so in this paper. Their model assumes that the counter ion atmospheres behave as stated by the Schwarz equations, an assumption of questionable validity as discussed above. They also assume that the bound water's dielectric behavior is characterized by a single relaxation time, instead of the spectrum reported by different investigators. But these are relatively minor points and their results support the general conclusions arrived at above, pointing out the highly frequency dependent fields applied to the various cellular compartments.

Some Nonthermal Interactions

We shall not be able to review the extensive literature on membrane excitation phenomena and some extraordinary membrane sensitivities. But a brief summary of biological manifestations of "pondermotoric" forces resulting from alternating electrical fields will be attempted (23, 24, 25). These forces emerge above a threshold field strength value E_{th}. Above this threshold the electric forces are larger than random thermal ones and the effect of interest emerges. A rather approximate threshold field strength value is given by the equation:

$$E_{th}\ R^{1.5} = 8 \cdot 10^{-5}\ \mathrm{Vcm}^{-\frac{1}{2}} \qquad (6)$$

with R particle radius (24). The equation applies fairly well for the phenomena of pearl chain formations, i.e. the alignment of suspended particles in the field direction. Other effects occur at comparable field values. As the equation indicates, the threshold value E_{th} is strongly dependent on radius R. For (R=1 μm E_{th} = 30 V/cm, but for cells of radius 100 μm E_{th} is only 0.03 V/cm. Previous theoretical contributions were largely based on energetic considerations (24,25,26). They have been criticized recently since the definition of the potential energy of a particle exposed to an alternating field remains evasive in the presence of dissipative and dispersive properties characteristic of biological structures (27). Sauer (27) provided a rigorous derivation for the force acting on a particle and the particle trajectories leading to the pearl chain alignment. Integration of this force over the path leading to particle alignment and equating this work with KT can be shown to yield a similar equation as the one stated in Equation 6, thus, perhaps explaining the apparent agreement of Equation 6 with experimental results.

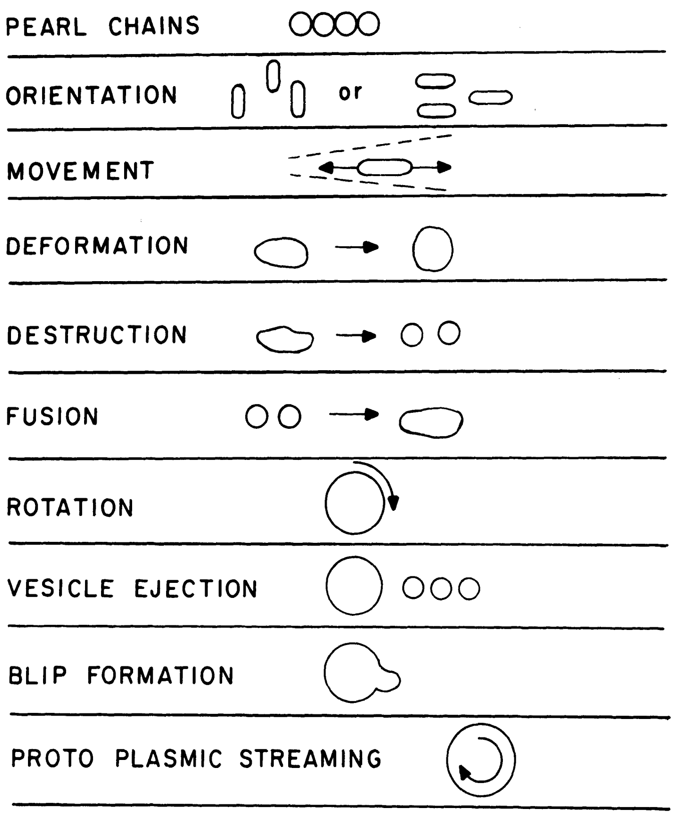


Figure 4. A summary of various field induced force effects acting on cells and depending on frequency and field strength.

Figure 4 attempts to summarize a variety of different phenomena which have been demonstrated so far and all appear to be caused by alternating field induced forces. These effects as well as the effects of fairly large potentials of the order of 1 V applied across membranes have been extensively studied in recent years. Zimmermann, Benz and their colleagues (28) have investigated in great detail the effects of briefly imposed membrane potentials and how they can be used to cause both reversible and irreversible local membrane destruction and how to utilize this effect to facilitate increased exchange of material from the extracellular to intracellular space. Neumann has recently demonstrated very large increases of cellular DNA uptake utilizing these electrical techniques (29). Of particular interest has been the possibility to greatly facilitate cellular fusion. Use is made of a combination of the pearl chain effect caused by moderate alternating field with brief strong pulses which result in membrane destruction at adjacent sites and formation of larger cells containing several nuclei. Using some of the principles outlined above, it should even be possible to selectively fuse the nuclei. An excellent and detailed recent summary of all of these efforts has been given by Zimmerman (28).

The relevance of these pondermotoric force effects to more subtle biological field effects is unclear. However, biological structures have a variety of means to amplify fairly small average in situ field strength values and current densities. For example, an average in situ field of the order of 1 V/cm is amplified below the cutoff frequency given by equation (2) into membrane potentials of some mV and membrane field strength values between 10 and 100 KV/cm. Average membrane current densities of only nA/cm2 across the membrane may be concentrated in membrane holes and gates and cellular gap junctions by orders of magnitude. Under such circumstances, it appears entirely possible that pondermotoric effects may play a role at local spots even though the average in situ fields are low. The investigation of such local events requires application of both topics discussed above, i.e. cellular coupling considerations as well as pondermotoric effects. In any case, the nonthermal effects discussed are likely to be frequency sensitive as demanded by the coupling considerations outlined first.

The effects discussed above have at first glance no apparent relationship to the modulation and window effects of primary interest at this meeting (30). Nor do they relate to the strong mm-resonances in the mm-wave range also discussed here (Keilmann, this volume). But they relate in a broader context, since they also demand highly frequency specific and for some compartments windowlike responses at least as far as frequency is concerned. Our discussions, furthermore, have been restricted to the linear level and not considered possible nonlinearities which may arise at certain cellular or subcellular locations.

REFERENCES

1. C.H. Durney, et al., Radio Frequency Radiation Dosimetry Handbook, 1st edition, September 1976, 2nd edition, May 1978, 3rd edition, August 1980. Reports SAM-TR-76-35/78-22 and 80-32, prepared for USAF School of Aeorspace Medicine, Brooks Air Force Base, Texas 78235. Additional material is contained in several articles of "Biological Effects and Medical Applications of Electromagnetic Energy", Special Issue of Proc. IEEE, 68, (1980).
2. H.P. Schwan, Biophysics of the Interaction of Electromagnetic Energy with Cells and Membranes in: "Biological Effects and Dosimetry of Nonionizing Radiation," M. Grandolfo, S.M. Michaelson and A. Rindi, (eds.) Plenum Publishing Corp. (1983).

3. H.P. Schwan, Electrical Properties of Tissue and Cell suspensions in "Advances in Biological and Medical Physics", J.H. Lawrence and C.A. Tobias, (eds.) Vol. V, p. 147, Academic Press, Inc., New York (1957).
4. G. Falk and P. Patt, Linear Electrical Properties of Striated Muscle Fibers Observed with Intracellular Electrodes, Proc. Royal Soc. B 160:69 (1964); P. Fatt, An Analysis of the Transverse Electrical Impedance of Striated Muscle, Proc. Roy. Soc. B 159:606 (1964).
5. H.P. Schwan, Electrical Properties of Cells: Principles, Some Recent Results and Some Unresolved Problems in: "The Biophysical Approach to Excitable Systems," W.S. Adelman and D. Goldman, (eds.) pp. 3-24, Plenum Press (1981).
6. E. Schifferdecker and E. Froemter, The AC Impedance of Necturus Gall Bladder Epithelium, Pflugers Arch., Eur. J. Physiol. 377:125-133 (1978).
7. G. Schwarz, A Theory of Low Freqency Dielectric Dispersion of Colloidal Particles in Electrolyte Solution, J. Phys. Chem. 66:2636 (1962).
8. R. Heinrich, M. Gaestel and R. Glaser, The Electric Potential Profile Across the Erythrocyte Membrane, J. Theor. Biol. 96:211-231 (1982).
9. H.P. Schwan, Electrical Properties of Bound Water, in: "Forms of Water in Biologic Systems", Annals of the New York Academy of Sciences, Vol. 125 (1965).
10. E.H. Grant, The Structure of Water Neighboring Proteins. Peptides and Amino Acids as Deduced from Dielectric Measurements, in: "Forms of Water in Biologic Systems", Annals of the New York Academy of Sciences, Vol. 125 (1965).
11. Oncley, J.L., Chem. Rev. 30:433 (1942).
12. S. Takashima and A. Minikata, Dielectric Behavior of Biological Macromolecules, in: "Digest of Dielectric Literature: National Research Council", Washington, D.C., Vol. 37, p. 602 (1975).
13. J.C. Maxwell, A Treatise on Electricity and Magnetism, Dover, New York, (1954).
14. H. Pauly and H.P. Schwan, The Impedance of a Suspension of Spherical Particles Surrounded by a Shell, ZS. f. Naturforschung, 14b:125-131 (1959).
15. H. Fricke, The Complex Conductivity of a Suspension of Stratified Particles of Spherical or Cylindrical Forms, J. Phys. Chem. 59:168 (1955).
16. H.P. Schwan and H.J. Morowitz, Electrical Properties of the Membranes of the Pleuropneumonia-like Organism A5969, Biophys. J. 3:395 (1962).
17. T. Hanai, T. Imakita and N. Koizumi, Analysis of Dielectric Relaxations of W/O Emulsions in the Light of Theories of Interfacial Polarization, Colloid Polymer Sci. 260:1029-1034, (1982).

18. A. Irimajiri, T. Hanai and A. Inouye, A Dielectric Theory of "Multi-Stratified Shell" Model with its Application to a Lymphoma Cell, J. Theor. Biol. 78:251-269 (1979).

19. H.P. Schwan, Interaction of Microwave and Radio Frequency Radiation with Biological Systems, IEEE Trans. on Microwave Theory Techniques, MTT-19:146-152 (1971).

20. H.P. Schwan, Effects of Microwave Radiation on Tissue -- A Survey of Basic Mechanisms, Non-Ionizing Radiation, Vol. 1, 1:23-31 (1969).

21. G.P. Drago and S. Ridella, Evaluation of Electrical Fields Inside a Biological Structure, Br. J. Cancer 45V:215 (1982).

22. G.P. Drago, M. Marchesi, S. Riddella and M Parodi, Electrical Modeling of Cells, presented at: "School on Biostructures in Erice 1982" to be published by Plenum Press.

23. H.P. Schwan and L.D. Sher, Alternating-Current Field-Induced Forces and Their Biological Implications, J. Electrochem. Soc., 116:170 (1969).

24. H.P. Schwan, Nonthermal Cellular Effects of Electromagnetic Fields: AC-Field Induced Pondermotoric Forces, Br. J. Cancer, 45V:220 (1982).

25. H.A. Pohl, Dielectrophoresis, The Behavior of matter in Nonuniform Electric Fields, London: Cambridge University Press (1978).

26. M. Saito and H.P. Schwan, The Time Constants of Pearl Chain Formation, in: "Biological Effects of Microwave Radiation", M.F. Peyton (ed.), Peyton, New York, Plenum Press, 1:85 (1961).

27. F.A. Sauer, Forces on Suspended Particles in the Electromagnetic Field in: "Coherent Excitations in Biological Systems," H. Frohlich (ed.), New York, Springer, in press (1983).

28. U. Zimmermann, Electric Field-Mediated Fusion and Related Electrical Phenomena, Biochimica et Biophysica Acta 694:227-277 (1982).

29. E. Neumann, M. Schaefer-Ridder, Y. Wang and P.H. Hofschneider, Gene Transfer into Mouse Lyoma Cells by Electroporation in High Electric Fields, The EMBO J., Vol. 1, No. 7:841-845 (1982).

30. W.R. Adey, Tissue Interactions With Nonionizing Electromagnetic Fields, Physiology Reviews, Vol. 61, No. 2 (1982).

ELECTROMAGNETIC SURFACE WAVE PENETRATION OF BIOLOGICAL MEDIA

Leonard S. Taylor

Depts. of Electrical Engineering and Radiation Oncology
University of Maryland
College Park, MD 20742

INTRODUCTION

The subject of this paper is a class of electromagnetic phenomena which is based upon the usual linear formulation of Maxwell's equations. It does, however, bear strongly upon the subject matter of this symposium devoted to non-linear electrodynamics. If we are concerned about the effects of electromagnetic fields upon living tissue, we must always begin with the question of the relationship between the externally applied field and the field which is acting upon the bio-molecules in the tissue. Thus the question of the penetration depth of applied fields is an important one. The results which we have obtained demonstrate that under certain special circumstances, at least, this depth may be considerably larger than usually assumed. Although these results appear to have important applications in the design of therapeutic devices, the question of their significance to the the problem of possible bioelectromagnetic effects is still a matter of speculation.

It is generally assumed that the penetration depth of an electromagnetic field in the radio frequency range or above into a lossy medium is limited by an exponential decay in which the e-folding distance is the electromagnetic "skin depth". At frequencies in the S-band portion of the microwave spectrum, two gigaherz or so, the electromagnetic skin depth is of the order of a centimeter in normal muscle and organ tissue. To be sure, some improvement in penetration of the induced thermal fields can be achieved by using surface cooling in combination with radiators arranged in some form of array. Nevertheless, it appears to be generally believed that the electromagnetic skin depth of a lossy material represents a physical property of the material which cannot be substantially circumvented by electromagnetic design techniques. Recently, however, we have been able to show[1] that this belief is not well-founded and that certain special electromagnetic modes can achieve greater penetration. This "discovery" was based on the observation that the usual expression for the electromagnetic skin depth is derived by considering the penetration of a homogeneous plane wave into a lossy medium. Homogeneous waves, however, represent only a portion (usually the most significant portion, to be sure) of the electromagnetic catalog. Inhomogeneous waves, which exist in the near fields of antennas, or as surface waves, which travel along the interfaces between media, are also a part of the electromagnetic wave type spectrum. Mathematically, these waves appear as a solution of Maxwell's equations when we assume that the wave propagation constant may be complex, even in a lossless medium. Physically, inhomogeneous waves possess a number of properties which differ from those of ordinary homogeneous waves; however, they are nonetheless "real" and play an important role in radio communications, for example. Daytime long-distance radio is made possible by the existence of the ground (Norton) wave, a surface wave which propagates along the earth/air interface, with amplitude decaying exponentially with height above the interface with a scale length of the order of tens of meters.

THE ZENNECK WAVE

Probably the earliest surface inhomogeneous wave type to be discovered is the famous Zenneck wave. One need only consider a planar interface (the x-y plane) with dielectric permittivity ε_1, for z>0 and ε_2 for z>0. The permittivities may be either real or complex. Setting the magnetic permeability equal to that of free space everywhere, and assuming no y-dependence for the field, it is easily shown that the magnetic intensities

$$H_y = H_0 e^{-\gamma_1 z - j\lambda x + j\omega t} \qquad z>0$$

$$H_y = H_0 e^{+\gamma_2 z - j\lambda x + j\omega t} \qquad z<0$$

will satisfy the wave equation and the boundary conditions at the interface (that is, H_y and the corresponding E_x are continuous across the interface), provided that

$$\gamma_1^2 = \frac{-\varepsilon_1^2}{\varepsilon_1 + \varepsilon_2} \mu_0 \omega^2$$

$$\gamma_2^2 = \frac{-\varepsilon_2^2}{\varepsilon_1 + \varepsilon_2} \mu_0 \omega^2$$

$$\lambda^2 = - \frac{\varepsilon_1 + \varepsilon_2}{\varepsilon_1 + \varepsilon_2} \mu_0 \omega^2$$

In the case which is usually considered, the earth/air interface, the permittivity in $z>0$ is that of free space and the region $z<0$ is lossy with $\varepsilon_2 = \varepsilon_w - j^{\sigma}{}_w/\omega$; $|\varepsilon_2|>>\varepsilon_1$. The wave in $z>0$ propagates in the x-direction and decays with height above $z=0$. Thus

$$\gamma_1{}^2 = - \varepsilon_0^2 \mu_0 \omega^2/\varepsilon_2$$

$$\lambda^2 = - \varepsilon_0 \mu_0 \omega^2$$

It is also readily seen that in this case the wave amplitude decays in $z<0$ with the usual electromagnetic skin depth.

The wave type described above is the Zenneck wave. It has had a rather unique role in the development of electromagnetic theory.[2] Briefly stated, the Zenneck wave was first described in 1907; as a result of an error in the paper published by Sommerfeld in 1909 on the radiation from a dipole above a lossy half-space, it was initially believed that the Zenneck wave was responsible for daytime long distance radio propagation. In 1917, Weyl obtained a different solution to the problem of the dipole above a lossy ground and a scientific controversy was generated that was not resolved until 1935, when Norton was able to find the error in Sommerfeld's calculation and demonstrated experimentally that the Zenneck wave was not the surface wave type that was involved in long-distance daytime propagation. For the next two decades the Zenneck wave maintained an anomalous existence in electromagnetic theory -- a wave type which could be simply demonstrated theoretically, but which had never been observed experimentally. The anomaly was finally resolved by Hill and Wait [3] in 1972; their results bear strongly upon the results to be described later in this paper. They found that a pure Zenneck wave is generated by an infinite antenna above an interface, but a finite antenna excites a wave which is Zenneck-like only in the near field and becomes the usual Norton ground wave within a distance that is so short for practical antennas that the Zenneck-like field is never observed.

PENETRATING WAVE FIELDS

We now consider a special case of the Zenneck wave which seems to have escaped previous notice. For $\varepsilon_1 >> |\varepsilon_2|$, the propagation constants given above reduce to the approximate forms

$$\gamma_1 \simeq j\sqrt{\varepsilon_1 \mu_0 \omega} \ , \quad \gamma_2 \simeq j\varepsilon_2 \sqrt{\mu_0 / \varepsilon_1 \omega}$$

and it may be seen that while the propagation factor in $z>0$ corresponds to the usual plane wave in a medium with permittivity ε_1, the wave field penetrates deeper into the medium in $z<0$. The penetration depth is given by the expression

$$\delta_w = \sqrt{\varepsilon_1 / \mu_0} / \sigma_w$$

Thus the electromagnetic skin depth has been increased by the factor $\sqrt{\varepsilon_1 |\varepsilon_2|}$ over the skin depth for an ordinary (homogeneous) plane wave.

Because of the capability of this type of field to penetrate deeper into the lossy medium, we have named waves with this property "pentrating waves". Certainly, this simple mathematical demonstration appears to show convincingly that it is physically possible for electromagnetic waves to propagate to relatively large depths in lossy media, and it appears very likely that other surface wave types with this same property also exist. However, in view of the results of Hill and Wait with regard to the limited extent of the ordinary Zenneck wave in the vicinity of a dipole above ground, it appears necessary to consider in detail the excitation of a penetrating field by a practical antenna. A problem of this type will be discussed in the next section.

MAGNETIC CURRENT LOOP ABOVE A LOSSY HALF-SPACE

In order to estimate the capabilites of a finite antenna configuration to excite a penetrating wave, we have calculated the field due to a magnetic ring current located in a space, $z>0$, above a half-space $z<0$ which is filled with a lossy dielectric. The region $z>0$ is assumed to have a very high permittivity, $\varepsilon_1 >> |\varepsilon_2|$. A magnetic ring source is equivalent to an electrically excited slot. Thus this example may be regarded as a simple model for an excited circumferential slot on a cylinder, or for a circular stripline patch antenna. The analysis follows well-established procedures in electomagnetic theory and leads to a Sommerfeld integral which must be evaluated by contour integration in the complex Fourier transform plane. The result depends upon the configuration of poles and zeroes which result from the particular choice of permittivities and ring current radius. We were able to cast the integral into a standard form in order to evaluate the z-dependence of the field in the lossy medium along the axis of the ring[1]. We found that the penetration depth of the field along the axis was increased, again by the factor $\sqrt{\varepsilon_1/\varepsilon_2}$ compared to the usual electromagnetic skin depth. However, an additional requirement was imposed -- the radius of the ring current must be large compared to the penetration depth. This requirement appears to be consonant with the results that show that the Zenneck wave only exists in the near field of finite antennas.

APPLICATIONS

Our mathematical results show that an electromagnetic field of the special type being discussed does exist in the vicinity of a finite antenna. By proper choice of dielectric materials we can obtain increased penetration depths into tissue media. For example, we could increase the penetration depth into organ and muscle tissue at 915MHz from about 2.3cm, the usual electromagnetic skin depth, to about 7cm using a wave launched in barium titanate or some equivalent high permittivity material. Of course, to match the model studied, some admittedly inconvenient arrangements might be required. The launcher would have to be quite large, of the order of a meter, and to match the planar geometry, if that is necessary, we would have to fill in with tissue-simulating dielectric material (bolus) which would have to be cooled. The figure illustrates this concept. Clearly, considering the improvement in hyperthermia treatment effectiveness which would result, the arrangement does offer some promise.

It is possible that these results have a wider practical significance in the context of bioelectromagnetic effects. It is known that surface waves are launched by temporally discontinuous fields. It has been shown theoretically [4] that pulsed sources launch Zenneck waves on dielectric interfaces. We can easily show that the particular limiting forms of the Zenneck wave that we have considered are also launched by pulsed sources. The question of the relative effectiveness of continous and pulsed sources in penetrating tissue has not been answered and may, conceivably, be a critical parameter in the problem of bioelectomagnetic interactions.

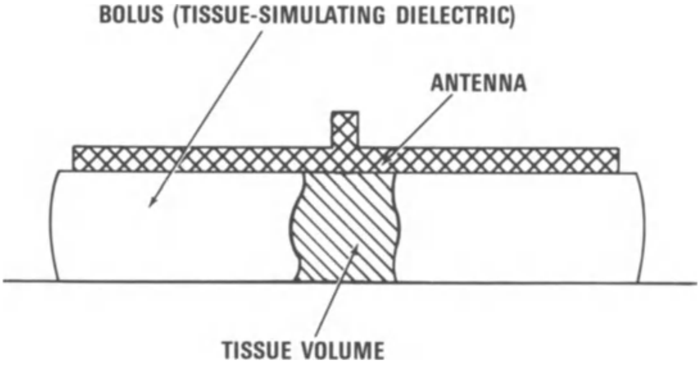


Fig. 1. Bolusing to obtain planar geometry.

REFERENCES

Taylor, L. S., Penetrating electromagnetic wave applicators, IEEE Trans. AP. To be published.

Banos, A., 1966, "Dipole Radiation in the Presence of a Conducting Half-Space", Pergamon Press, Oxford.

Hill, D. A., and Wait, J. R., 1972, Excitation of the Zenneck surface wave by a vertical aperture, Radio Science, 13: 969.

Van der Pol, B., 1956, On discontinuous electromagnetic waves and the occurrence of a surface wave. IRE Trans. AP 4:288.

MATHEMATICAL AND PHYSICAL MODELS OF NONLINEAR PROCESSES

CONCEPTUAL MODELS OF CO-OPERATIVE PROCESSES IN NONLINEAR ELECTRODYNAMICS

R. K. Bullough

Department of Mathematics
UMIST, P.O. Box 88
Manchester M60 1QD, U.K.

INTRODUCTION

This paper surveys conceptual models for the interaction of radiation and matter at small energies ($\lesssim$ 100 e.v.) but up to relatively large intensities ($\sim$ 100 W cm.$^{-2}$). My idea in presenting this survey at this Research Conference on Non-linear Electrodynamics in Biological Systems is two-fold: first of all it seems necessary to review non-linear electrodynamics itself in order to see how the non-linearities arise and what their consequences are; secondly, by covering this *energy* range we cover the electromagnetic (e.m.) spectrum into the far ultra-violet which covers the atomic processes likely to be relevant to biology, whilst by covering this *intensity* range we can cover all the collective processes, static or dynamic, which have been found in the studies of quantum optics and which *may* be applicable in biology. These collective processes involve dynamical phase transitions far from equilibrium, distinctions between coherent and incoherent processes, collective coherent energy transmission without energy loss, collective radiation damping processes which markedly accelerate that damping, quenching of damping, optical amplification, evolution of coherent e.m. fields through incoherent pumping, coherent field induced rapid switching between coherent and incoherent states, and other processes I shall mention. All of these rest on the non-linearities in the e.m. interactions. I therefore start by introducing these.

2. THE ELECTROMAGNETIC NON-LINEARITIES

In quantised form, namely as quantum electrodynamics, current theories of the interaction of radiation and matter are successful theories in so far as numbers can be computed from them which exceed in precision the accuracy of present experimental observations with which they already agree: the numbers usually quoted in this connection are the radiative level shifts (e.g. the Lamb-Rutherford shift in Hydrogen) or the anomalous moment of the electron. The striking fact about q.e.d. in the present connection is that it is a strictly linear theory in so far as it quantises e.m. fields which satisfy the linear Maxwell e.m. field equations. Quantum chromodynamics which involves Yang-Mills fields analogous to the e.m. field is actually non-linear because the Yang-Mills fields involve SU_2 or other gauge groups; but the gauge group of the e.m. field is the commutative unitary group U(1) and the e.m. field is linear in consequence.

The nonlinear e.m. interactions arise in q.e.d. as follows: matter, including biological matter, is made up of atoms, or molecules or ions and electrons. Radiation falling on one particular atom or electron inside a material body is made up of externally applied radiation plus the collective effects of radiation scattered from all the other atoms or electrons. It is this collective radiation from all the other atoms which is the source of the non-linearity. We use nonlinearity in its mathematical sense: the equations become nonlinear. As we shall see, an external field can drive a single atom linearly or nonlinearly in the field amplitude, but the equations of motion themselves remain linear. Obviously, such nonlinearities are important only when the fields are strong. This condition implies the number of atoms must be large to obtain mathematical nonlinearity of the collective equations.

Of course this is not the only source of non-linearity: electrons and ions interact through their Coulomb potentials, that is through the 'longitudinal' part of the e.m. field; radiation concerns the 'transverse' part. But in a crystal lattice, for example, or a DNA molecule for that matter, electrons may interact through Coulomb fields with displacement of ions about their equilibrium configuration, that is to the 'phonons'; and this may induce nonlinearities. An example is the Su - Schrieffer - Heeger model[1] applied to the excitations of trans poly-acetylene (described by both W. P. Su[2] and A. R. Bishop[2] at this meeting) and there are of course many other such solid state models. In this paper we concentrate on the non-linearities in the radiation field induced by collective transverse fields falling on particular atoms.

A model which has had profound effects on the evolution of the theory of quantum optics in the last 25 years is a model of two-level atoms ('quantum optics' covers amongst other things the theory of the laser or maser, a system very dependent on a non-linearity as we show in §4). A single two-level atom is precisely that — namely two non-degenerate atomic levels with an energy separation we will call $\hbar\omega_s$, so that ω_s is the resonance frequency. The states are $|s>$ and $|o>$; the eigen energies are E_s and E_o; and $E_s - E_o = \hbar\omega_s$.

The Hamiltonian operator can therefore be taken in the form $H_o = E_s\,|s><s| + E_o\,|o><o|$; for with $<s|o> = 0$, the eigenstates are $|s>$ and $|o>$ and the eigen energies are E_s, E_o. It is convenient to map this Hamiltonian on to one for a spin-$\frac{1}{2}$ system. Put

$$H_o = \tfrac{1}{2}\hbar\omega_s\,\sigma_z \tag{2.1}$$

where σ_z is the Pauli spin matrix $\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$: then $|s>$, energy $+\frac{1}{2}\hbar\omega_s$, is 'spin up' and $|o>$, energy $-\frac{1}{2}\hbar\omega_s$ is 'spin down'.

There will also be a matrix element for the transitions $|o> \to |s>$ or $|s> \to |o>$: we work in dipole approximation so the relevant matrix element is $<s|\underset{\sim}{p}|o>$, where $\underset{\sim}{p}$ is the dipole operator. We shall write $p\hat{\underset{\sim}{u}}$, magnitude p, for this matrix element.

In semi-classical radiation theory the e.m. field is a c-number satisfying Maxwell's equations. H_0, the total Hamiltonian, H. will include the interaction

$$H_{int} = -\underset{\sim}{p}\cdot\underset{\sim}{E}(t) \tag{2.2}$$

where $\underset{\sim}{E}(t)$ is the electric field falling on the atom at time t (and is a c-number). A little thought will show that $H_{int} = -p\hat{\underset{\sim}{u}}\cdot\underset{\sim}{E}(t)\sigma_x$ where σ_x is the Pauli matrix $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$. Then from Heisenberg's equation of motion and the commutation relations for Pauli matrices we find (with $E(t) \equiv \hat{\underset{\sim}{u}}\cdot\underset{\sim}{E}(t)$)

$$\begin{aligned}\dot\sigma_x &= -\omega_s\sigma_y\\ \dot\sigma_y &= +\omega_s\sigma_x + 2p\hbar^{-1}E(t)\sigma_z\\ \dot\sigma_z &= -2p\hbar^{-1}E(t)\sigma_y \quad .\end{aligned} \tag{2.3}$$

This is the operator form of the 'Bloch' equation — the name being taken from the theory of spin resonsnce. Equivalently

$$\dot{\underset{\sim}{\sigma}} = \underset{\sim}{\omega}\times\underset{\sim}{\sigma}\ ;\ \underset{\sim}{\omega} = (-2p\hbar^{-1}E(t), 0, \omega_s)\ . \tag{2.4}$$

Note that $\underset{\sim}{\sigma} = \{\sigma_x, \sigma_y, \sigma_z\}$ depends on t as a Heisenberg operator: $\underset{\sim}{\sigma}(t) = i\hbar^{-1}[H,\underset{\sim}{\sigma}]$ with $H = H_0 + H_{int}$, $\underset{\sim}{\sigma} = d\sigma/dt$. Equation (2.4) may be interpreted as describing a spin. The rate of charge of angular momentum depends on the couple applied, therefore this couple depends linearly on the "magnetic moment" and hence linearly on the spin $\underset{\sim}{\sigma}$ itself. Note that equation (2.4) is in the Heisenberg representation in which operators, $\underset{\sim}{\sigma}$, evolve in time. It is precisely equivalent ot Schrodinger's equation for a two-level atom with Hamiltonian H^3.

Notice that (2.3) or (2.4) is a strictly linear operator system. The external field $E(t)$ is prescribed and does not depend on the motion of the $\underset{\sim}{\sigma}(t)$. However, this is only true in the single atom case. In the many atom case the bilinear combinations $E(t)\sigma_z(t)$ and $E(t)\sigma_y(t)$ driving $\dot{\sigma}_y$ and $\dot{\sigma}_z$ respectively become strict non-linearities. This happens in the following way.

In the many-atom case $\underset{\sim}{\sigma}$ must depend on the position $\underset{\sim}{x}$ occupied by the atom, as well as on time t. Likewise $E(t)$ is to be evaluated at $\underset{\sim}{x}$ and depends on $\underset{\sim}{x}$, and t. It is therefore determined by Maxwell's wave equation

$$\nabla^2\underset{\sim}{E}(\underset{\sim}{x},t) - c^{-2}\partial^2\underset{\sim}{E}(\underset{\sim}{x},t)/\partial t^2 = 4\pi n\partial^2\underset{\sim}{P}(\underset{\sim}{x},t)/\partial t^2 \tag{2.5}$$

where $n\underset{\sim}{P}(\underset{\sim}{x},t)$ is the dipole density at $(\underset{\sim}{x},t)$. We suppose a uniform number density n of atoms, $\underset{\sim}{P}(\underset{\sim}{x},t)$ is scaled to a single atom at $\underset{\sim}{x}$. But plainly

$$\underset{\sim}{P}(\underset{\sim}{x},t) = p\hat{\underset{\sim}{u}}<\sigma_x(\underset{\sim}{x},t)> \tag{2.6}$$

where $<\cdots>$ denotes quantum mechanical expectation value (of the operator density $\sigma_x(\underset{\sim}{x},t)$). Thus $E(t)$ in (2.3), evaluated at $\underset{\sim}{x}$, is driven by $<\sigma_x(\underset{\sim}{x},t)>$: so $E(\underset{\sim}{x},t)$ is a linear functional of $<\sigma_x(\underset{\sim}{x},t)>$ and the products $\sigma_z(\underset{\sim}{x},t)E(\underset{\sim}{x},t)$ and $\sigma_y(\underset{\sim}{x},t)E(\underset{\sim}{x},t)$ are bilinear (that is non-linear) functionals of $\underset{\sim}{\sigma}(\underset{\sim}{x},t)$. This is how the collective non-linearity arises.

Equations (2.3) and (2.5) together constitute the Bloch-Maxwell (B-M) system of equations. Typically (in semi-classical theory) one takes the expectation value of (2.3) and defines the Bloch vector density $\underset{\sim}{r}(\underset{\sim}{x},t) \equiv <\underset{\sim}{\sigma}(x,t)>$. The coupled B-M system is then

$$\dot{\underset{\sim}{r}}(\underset{\sim}{x},t) = \underset{\sim}{\omega}(\underset{\sim}{x},t) \times \underset{\sim}{r}(\underset{\sim}{x},t) \; ; \; \underset{\sim}{\omega}(\underset{\sim}{x},t) = (-2p\hbar^{-1}E(\underset{\sim}{x},t),0,\omega_s) \tag{2.7a}$$

$$\nabla^2E(\underset{\sim}{x},t) - c^{-2}\partial^2E(\underset{\sim}{x},t)/\partial t^2 = 4\pi np\partial^2\underset{\sim}{r}_x(\underset{\sim}{x},t)/\partial t^2 \tag{2.7b}$$

where $\underset{\sim}{r}(\underset{\sim}{x},t) = (r_x(\underset{\sim}{x},t),\ r_y(\underset{\sim}{x},t),\ r_z(\underset{\sim}{x},t))$. In particular

$$r_z(\underset{\sim}{x},t) \equiv <\sigma_z(\underset{\sim}{x},t)> = <\Psi(\underset{\sim}{x},t)|\sigma_z(\underset{\sim}{x},0)|\Psi(\underset{\sim}{x},t)> \tag{2.8a}$$

where $|\Psi(x,t)>$ is in Schrödinger representation. Then if

$$|\Psi(\underset{\sim}{x},t)> = C_s(\underset{\sim}{x},t)|s> + C_o(x,t)|o> \quad , \tag{2.8b}$$

$$r_z(x,t) = C_s(\underset{\sim}{x},t)C_s^*(\underset{\sim}{x},t) - C_o(\underset{\sim}{x},t)C_o^*(\underset{\sim}{x},t) \quad . \tag{2.8c}$$

This is just the inversion density per atom at $\underset{\sim}{x}$ at time t. It is expressible in terms of the density matrix $\rho_{\alpha\beta}(\underset{\sim}{x},t)$ at $(\underset{\sim}{x},t)$ by

$r_z(\underset{\sim}{x},t) = \rho_{ss}(\underset{\sim}{x},t) - \rho_{oo}(\underset{\sim}{x},t)$. It tells us what density $n\rho_{ss}(\underset{\sim}{x},t)$ of atoms is in the upper state s at $\underset{\sim}{x}$ at time t, etc. Equation (2.7a) also means, as with (213) and (2.4), that r^2 is a constant of the motion: $\partial \underset{\sim}{r}^2(\underset{\sim}{x},t)/\partial t = 0$. Moreover $\underset{\sim}{r}^2 = (|C_s|^2+|C_o|^2)^2$ so its constant value is 1. This enables us to define the Bloch sphere if there is no damping.

The elementary approach to the B-M equations is to solve them for *weak* fields E(x,t). Then $r_z(\underset{\sim}{x},t)$ is approximately a constant of the motion (to $O(|E|)^2$). Suppose $r_z(\underset{\sim}{x},0) = -1$ for all $\underset{\sim}{x}$, i.e. all atoms are in their ground states at $t=0$. Then $r_z(\underset{\sim}{x},0) = -1$ for all t (to $O(|\underset{\sim}{E}|)^2$ and the Bloch equation is

$$\dot{r}_x(\underset{\sim}{x},t) = -\omega_s r_y(\underset{\sim}{x},t)$$

$$\dot{r}_y(\underset{\sim}{x},t) = +\omega_s r_x(\underset{\sim}{x},t) - 2p\hbar^{-1}E(\underset{\sim}{x},t) \tag{2.9a}$$

so

$$\dot{r}_x(\underset{\sim}{x},t) + \omega_s^2 r_x(\underset{\sim}{x},t) = 2p\hbar^{-1}\omega_s \underset{\sim}{E}(x,t) \quad . \tag{2.9b}$$

This is just the driven harmonic oscillator.

Suppose now

$$E(\underset{\sim}{x},t) = E_o e^{-i(\omega t-kz)} \quad . \tag{2.10}$$

Then

$$pr_x(\underset{\sim}{x},t) = \frac{2p^2\hbar^{-1}\omega_s E(\underset{\sim}{x},t)}{\omega_s^2 - \omega^2} \equiv \alpha(\omega)E(\underset{\sim}{x},t) \tag{2.11}$$

and the Maxwell equation requires

$$-k^2 + \omega^2 c^{-2} = -4\pi n\alpha(\omega)\omega^2 c^{-2} \quad . \tag{2.12a}$$

So k gains the refractive index $m(\omega)$ i.e. $k = m(\omega)\omega c^{-1}$, and

$$m^2(\omega) - 1 = 4\pi n\alpha(\omega) \tag{2.12b}$$

in which $\alpha(\omega)$ is the usual Kramers-Heisenberg polarisability (for a 2-level atom).

It is now plain the non-linearity comes in through the inversion density $r_z(\underset{\sim}{x},t)$. If the fields are stronger it will evolve and the B-M system becomes non-linear as indicated above. This explains why the 2-level atom can be a good model even though all atoms have many levels. One can easily check that unless $\omega \approx \omega_s$, i.e. the field is almost resonant with the 2-level transition, $r_z(\underset{\sim}{x},t)$ remains a constant of the motion. This applies to all non-resonant transitions in a multi-level atom. Thus we can deal with resonant (or almost resonant) optical phenomena and isolate the one atomic transition that is (almost) resonant. The remaining levels provide a linear background which is not interesting. If the levels are degenerate or almost degenerate we have to include additional Bloch type equations to describe the evolution of these. The example of degenerate self-induced transparency (SIT, see §3) is treated in Ref. [4].

Notice we already have one new result: if the atoms lie in their upper states at $t=0$, then $r_z(\underset{\sim}{x},t)=+1$ for all $\underset{\sim}{x}$; and $\alpha(\omega)\to -\alpha(\omega)$, so that, for $\omega<\omega_s$, $m^2(\omega)<1$ and harmonic modes (2.10) travel at speeds $>c$. This model is the simplest form of amplifier: the model with $r_z(\underset{\sim}{x},0)=-1$ for all $\underset{\sim}{x}$ is the attenuator[3]. Both of these models become more interesting when the non-linearity is allowed to act. In the examples which follow the non-linearity is essential.

3. SELF-INDUCED TRANSPARENCY

Self-induced transparency (SIT) is observed as short (10^{-9}sec) coherent optical pulses which enter a resonant medium and reshape to characteristic hyperbolic secant (sech) shaped forms and re-emerge without further loss of energy. In principle e.m. energy is transmitted without loss providing the coherent pulse has the appropriate sech shape. These sech pulses have characteristic 'areas" -- namely 2π, after proper scaling. A pulse of area θ in $(2n+1)\pi>\theta>(2n-1)\pi$ actually evolves and breaks up into a train of n 2π-pulses each of which transmit without energy loss. The observed delay can be a striking feature -- as though the pulse, which is a pulse envelope, gains a refractive index perhaps as big as 10^3 even in experiments on metal vapour where n ~10^{11} or 10^{12} atoms $cc.^{-1}$(very small values even compared with those for most gases).

The coupled B-M system (2.7) describes all this in the following way: approximations are introduced in order to deal with the non-linearity. These rely on the existence of two time-scales in the problem. One is the optical frequency $\omega\approx\omega_s\sim10^{15}$ Hz —— time scale 10^{-15} sec. The other is the pulse life time ~ 10^{-9} sec. This life time is controlled by the Rabi frequency $\Omega\equiv pE_o\hbar^{-1}$ where E_o is the pulse amplitude: if $p\approx5$ Debyes (5×10^{-18} c.g.s. units) and E_o corresponds to Wcm^{-2} power at peak, then $\Omega\sim10^9$ Hz. The ratio of these two time scales is ~ 10^6: hence, if E_o varies slowly (on the 10^{-9} sec. time scale), we can expect to neglect terms $O(\Omega\omega_s^{-1})$.

One sets

$$E(z,t) = \hbar p^{-1}\varepsilon(z,t)\cos\Phi(z,t)$$

$$\Phi(z,t) = kz-\omega t+\phi(z,t) \quad ; \quad \omega=ck$$

$$r_x(z,t) = Q(z,t)\cos\Phi + P(z,t)\sin\Phi \tag{3.1}$$

and $\omega \approx \omega_s$, $k \approx k_s = \omega_s c^{-1}$. Then npQ and npP are the in-phase and phase quadrature components of the dipole moment density. These vary on the scale Ω^{-1} in time. By assumption $\varepsilon(z,t)$, P and Q vary slowly in space z also, $(\partial\varepsilon/\partial z)k_s \ll 1$, etc. Notice that we look for plane wave solutions travelling along z alone. In practice, pulses have finite transverse apertures and self-focusing occurs so the details become complicated[5]. Nevertheless, the one-dimensional behaviour depending on z alone is qualitatively correct. Note also that we include a slowly varying phase $\phi(z,t)$ varying on time and space scales so that $(\partial\phi/\partial t)\omega_s \ll 1$, $(\partial\phi/\partial z)k_s \ll 1$.

For completeness of notation we re-define

$$r_z(z,t) \equiv N(z,t) \quad . \tag{3.2}$$

We also extend the definitions of P, Q, N so that they depend explicitly on the 'detuning' $\Delta\omega' \equiv \omega_s' - \omega$, where ω_s' is introduced for ω_s now for the reason that the atoms are supposed to resonate over a range of values ω_s' with distribution $g(\Delta\omega')$. In vapours this range arises through Doppler broadening; in solids local interactions spread the resonance. This broadening of the resonance is called 'inhomogeneous broadening'.

With the model extended in this respect we reach the equations (3.3a-c) — simply by inserting (3.1) and (3.2) into the B-M equations (2.7) and equating coefficients of sin Φ and cos Φ. (This is an acceptable procedure if no other harmonic modes sin Φ' and cos Φ' are close on the time scale Ω^{-1}.) The equations are

$$\begin{aligned} \partial\varepsilon/\partial z + c^{-1}\partial\varepsilon/\partial t &= \alpha \langle P(z,t;\ \Delta\omega')\rangle \\ \varepsilon(z,t)\ (\partial\phi/\partial z + c^{-1}\partial\phi/\partial t) &= -\alpha \langle Q(z,t;\ \Delta\omega')\rangle \end{aligned} \tag{3.3a}$$

$$\begin{aligned} \partial Q/\partial t &= -(\Delta\omega' + \partial\phi/\partial t)\ P \\ \partial P/\partial t &= \varepsilon N + (\Delta\omega' + \partial\phi/\partial t)\ Q \\ \partial N/\partial t &= -\varepsilon P \end{aligned} \tag{3.3b}$$

Temporarily

$$\langle f(\Delta\omega')\rangle \equiv \int_{-\infty}^{\infty} g(\Delta\omega')\ f(\Delta\omega')\ d(\Delta\omega') \tag{3.3c}$$

and is not a quantum expectation value, and

$$\alpha c \equiv 2\pi n p^2 \omega_s \hbar^{-1} \equiv \tau_c^{-2} \quad . \tag{3.4}$$

The time τ_c defines an important time scale: it is the 'maximum co-operation time'[5,6] and relates to two other times τ_E and τ_{SF}, $\tau_E\tau_{SF} = \tau_c^2$, and the times τ_E and τ_{SF} are respectively the photon escape time $2Lc^{-1}$, and the 'superfluorescence' time. The length of the atomic sample (along z) is L, and $\tau_{SF} = (4\pi np^2\omega_s c^{-1}\hbar L^{-1})^{-1}$ is the time scale of co-operative i.e. collective decay processes in the optical system. It is characterised by a dependence on n and hence on the number of atoms N in the sample. Large N means small τ_{SF} and fast collective decay. We mention this again briefly in §6 in connection with 'super-radiance'. Super-radiance concerns collective decay inside a cavity where the geometry of the process is predetermined. Superfluorescence is similar collective (co-operative) decay but the process forms its own geometry for the direction of actual emission[5,6,7].

The equations (3.3) prove to be an 'integrable system' in the technical sense[5,8]. They have 'soliton' solutions[5,8] and it is these which are the optical pulses and propagate without loss. This is not the place to study integrable systems and the reader is referred to our book[8], for example. Nor is this the place to treat (3.3) in depth. We first make several remarks and then illustrate the solutions of (3.3) in a particular simplified form.

The first remark is that equations (3.3a) are the Maxwell equations but now in slowly varying envelope and phase approximation (SVEPA). The equations (3.3b) are the Bloch equations in SVEPA. In reaching these we make not only the SVEPA, but also the rotating wave approximation (r.w.a.): this neglects all high frequency oscillations at $2\times\omega$ ($\approx 2\times\omega_s$) and $3\times\omega$ ($3\times\omega_s$), etc. Notice that (3.3b) has Bloch equation *form* very like (2.3); but it refers to envelopes (and frequency differences $\Delta\omega'$).

The second remark, *en passant*, is that the B-M equations (2.7) can be reduced to the 'Reduced Maxwell Bloch (RMB)' equations simply by looking for solutions travelling along +z alone[5]. The Maxwell equation becomes $\partial E/\partial z + c^{-1}\partial E/\partial t = -2\pi np\,\partial r_x/\partial t$ (simply) and is valid for small n. The Bloch equation is not changed. The RMB equations are integrable in the sense discussed but the M-B equations themselves are not.

The third remark is that equations (3.3) are integrable even with inhomogeneous broadening included. To see what this means consider the following: each atom has its own Bloch vector $\underset{\sim}{r}(t)$ (i.e. atoms at $\underset{\sim}{x}$ have $\underset{\sim}{r}(\underset{\sim}{x},t)$). It is easily checked that a solution of $\dot{\underset{\sim}{r}}(t) = \underset{\sim}{\omega}\times\underset{\sim}{r}(t)$ is $E(t) = 0$, $r_z(t) = \cos\theta$, $r_x = \sin\theta\cos\omega_s t$, $r_y = \sin\theta\sin\omega_s t$, where $\theta = \text{constant}$. Hence the Bloch vector precesses at angular velocity ω_s about the vector's z-axis (N.B. this is not the co-ordinate z-axis: the Bloch sphere is its own separate space). Then $\underset{\sim}{r}$'s with resonances ω_s' precess at ω_s' and get out of phase with each other. The coherence expressed by the collective action of the

Bloch vectors $\underset{\sim}{r}$ is lost. A remarkable feature of the SIT pulses which solve (3.3) is that this dephasing is quenched by the action of the pulses. In this way coherent information propagates without loss in the medium.

A fourth remark is then that propagation without loss of energy or coherence in biological materials is the key idea behind Davydov's treatment[9,10] of energy and information transfer along the protein α-helix: Davydov's equations, which relate to the integrable 'Non-linear Schrödinger equation'[8] are not integrable themselves (as far as I know) and his excitation is a 'solitary wave'. Solitary waves also propagate without distortion and loss[8,11]. Interpreted in this very much broader sense the soliton has begun to play a very significant conceptual role in biology and is one reason for this interdisciplinary conference (see particularly the papers by Scott concerning α-helix transmission and by Lawrence and Adey concerning soliton propagation across cell membranes). The other idea of equal significance to biology at the present time is Fröhlich's oscillator[13]. We describe a number of non-linear oscillators derived from versions of the equations (3.3) later (§§4 and 5). It is interesting to speculate what mathematical connections, if any, we can ultimately establish between these different ideas.

We now return to (3.3) and illustrate its behavior in the simplified case of 'sharp line' resonant solutions, i.e. $g(\Delta\omega') = \delta(\Delta\omega')$. The Bloch equations are then

$$\partial Q/\partial t = 0, \quad \partial P/\partial t = \varepsilon N, \quad \partial N/\partial T = -\varepsilon P \tag{3.5}$$

solved by $Q \equiv 0$, $P = -\sin\sigma$, $N = -\cos\sigma$; $\partial\sigma/\partial t = \varepsilon$ (σ is now just a function -- nothing to do with Pauli matrices!) Set $\dot{\sigma} = \varepsilon$ in (3.3a) and solve for $\phi = 0$ (no 'chirps'). Then substitute

$$c\,\frac{\partial^2\sigma}{\partial t\,\partial z} + \frac{\partial^2\sigma}{\partial t^2} = -\alpha\sin\sigma \quad . \tag{3.6a}$$

If $\sqrt{\alpha}\,ct = \tau\sqrt{c}$, $\sqrt{\alpha}\,(ct - 2z) = \xi\sqrt{c}$

$$\frac{\partial^2\sigma}{\partial\xi^2} - \frac{\partial^2\sigma}{\partial\tau^2} = \sin\sigma \tag{3.6b}$$

and this is the now famous 'sine-Gordon equation'[8]

$$\frac{\partial^2 u}{\partial z^2} - \frac{1}{c^2}\,\frac{\partial^2\sigma}{\partial\tau^2} = m^2\sin u \tag{3.7}$$

with mass m and speed c set equal to unity by the scaling.

The sine-Gordon equation is an integrable system and arises in many physical contexts: Josephson junctions, spin waves in liquid ^{3}He (A-phase), crystal dislocation theory[11], spin waves in one dimensional ferromagnetic and antiferromagnetic crystals ($CsNiFe_3$, TMMC for example[14]). It is hard to believe that the sine-Gordon equation will not play some fundamental role in biological systems where coherent transmissions of information occur.

We quote here the kink solution of the s-G equation (3.7) namely

$$u(z,t) = 4\tan^{-1}\exp\left\{\frac{m(z-Vc^{-1}t)}{\sqrt{1-V^2c^{-2}}}\right\} . \tag{3.8}$$

This kink solution yields an optical SIT pulse through (3.6) which is

$$\sigma(z,t) = -4\tan^{-1}\exp\left\{\frac{(z-Ut)}{\sqrt{U(1-U)}}\right\} \tag{3.9a}$$

where $c=1$ and $U=\frac{1}{2}(c+V)$; and since $\varepsilon(z,t)=\partial\sigma/\partial t$,

$$\varepsilon(z,t) = 2\tau_p^{-1}\,\mathrm{sech}\,\{\tau_p^{-1}(t-U^{-1}z)\} \tag{3.9b}$$

where the pulse life time $\tau_p=\hbar/pE_o=\Omega^{-1}$, and $cU^{-1}-1=4\alpha c\tau_p^2$. Notice how the speed U of the pulse (3.9b) in the original co-ordinates depends on the pulse amplitude E_o. Notice too that the 'area' $\Theta=\int_{-\infty}^{\infty}(\partial\sigma/\partial t)dt$ of this SIT pulse is precisely 2π. An arbitrary pulse breaks up into a train of kinks (area 2π) or antikinks (area -2π) and a train of SIT pulses made up of $\pm 2\pi$ pulses emerges from the single incoming pulse. These have been seen very clearly (in ^{87}Rb vapour[15], for example). In principle there could be breather pulses too but we do not discuss breathers here (Ref.[8]).

The theory of SIT is a non-linear theory of the attenuator (ref. §2). The corresponding theory for the amplifier (in this case for degenerate 2-level atoms) has just been reported[16].

4. EFFECTS OF RADIATION OR OTHER DAMPING

The radiation damping time T_2 is equal to $\gamma_o^{-1}\equiv 2\Gamma_o^{-1}$ where (for two-level atoms) Γ_o is the Einstein A-coefficient $\frac{4}{3}p^2\omega_s^3c^{-3}\hbar^{-1}$. For $\omega_s\sim 10^{15}$ Hz., $\Gamma_o\sim 10^8$ Hz. Thus if $\Omega\sim 10^9$ Hz, the SIT pulse can act, moving the atomic Bloch vectors through complete twists of 2π (ground state back to ground state) without the radiative damping causing much change. An analysis valid for long times $\gg\gamma_o^{-1}$ is given in Ref. 17. However, if the driving field is not pulsed, but is for example a steady oscillation like (2.10), radiation damping becomes extremely important and changes the character of the solutions. The same is true for any damping process described by T_2 not just radiation damping.

If (2.10) is a solution, then since $\Omega=pE_o\hbar^{-1}$, the Bloch equation in the envelope form (3.3b) becomes

$$\frac{\partial Q}{\partial t} = -\Delta\omega P - T_2^{-1}Q$$

$$\frac{\partial P}{\partial t} = \Omega N + \Delta\omega Q - T_2^{-1}P$$

$$\frac{\partial N}{\partial t} = T_1^{-1}(\Lambda - N) - \Omega P \tag{4.1}$$

where $T_1 = \Gamma_o^{-1}$, $T_2 = 2\Gamma_o^{-1} = 2T_1$, and $\Delta\omega = \omega_s - \omega$. When $\Omega = 0$, the solution is $P = Q = 0$ (no dipoles) and $N = \Lambda = -1$, (ground-state atoms), where Λ is a new parameter describing 'pumping' of the system. It is incoherent pumping with the property that N settles on $+\Lambda$, whilst all coherence (solely in the phase of the dipole terms P,Q) is lost. Pumping may be incoherent in this form, or it can be coherent pumping through E (i.e. through Ω) $\neq 0$. We include coherent pumping in equation (4.11) below.

Consider the case of sharp lines and exact resonance $\Delta\omega = 0$. There is a steady state solution of (4.1), namely

$$Q = 0,\ P = T_2\Omega N,\ N = \Lambda[1 + \Omega^2 T_1 T_2]^{-1} \ . \tag{4.2}$$

Note that $1 \geq \Lambda \geq -1$ and no incoherent pumping at all corresponds to $\Lambda = -1$: *positive* pumping has $1 \geq \Lambda > 0$. Notice also that P is an envelope so it is now the constant amplitude of a plane wave solution $Pe^{-i(\omega t - kz)}$ of the original M-B equations. It cannot satisfy the phase and envelope Maxwell equations (3.3a). It is usual to ignore chirps (set $\phi \equiv 0$) and recognize the original Maxwell equation (2.7b) is damped by some (empirical) damping constant κ so that the envelopes satisfy

$$\frac{\partial \varepsilon}{\partial z} + c^{-1}\frac{\partial \varepsilon}{\partial t} + \kappa\varepsilon = \alpha P \quad . \tag{4.3}$$

Consequently the steady solution is now found from

$$\frac{\partial \varepsilon}{\partial z} = \alpha\kappa^{-1}T_2\Omega\Lambda[1 + \Omega^2 T_1 T_2]^{-1} \ , \quad \Omega = \frac{\partial \varepsilon}{\partial z} \tag{4.4}$$

There are several interesting features about this solution. Let us suppose spatial variations are unimportant. This can be so either because the atoms lie inside a cavity, as in the laser, and the cavity has its natural modes (take the 'single mode' laser), or because 'on average' the spatial variation is unimportant. If we integrate (4.3) over $\int_0^L dz$ and divide by L and do the same for P and Q, we obtain a so called 'mean field' theory[18]. In general, spatial variation or mode coupling is important so that we can only get a qualitative theory, at best, this way. In any case ΩN, (that is εN), in (4.1) has to be averaged as $L^{-1}\int_0^L \varepsilon dz \ \ L^{-1}\int_0^L N dz$, which is possibly a serious approximation.

Still once spatial variation is unimportant (4.3) becomes

$$c^{-1}\frac{\partial \varepsilon}{\partial t} + \kappa\varepsilon = \alpha P \tag{4.5}$$

so (4.4) can be extended to

$$c^{-1}\frac{\partial\varepsilon}{\partial t} = -\kappa\varepsilon\,\{1-4C\Lambda[1+\varepsilon^2 T_1 T_2]^{-1}\} \tag{4.6}$$

with the 'co-operation number' C defined by

$$C = \frac{1}{4}\alpha\kappa^{-1}T_2 \quad ; \tag{4.7}$$

$C > 0$, and is a pure number. Notice that (4.6) is not steady in ε, but P,N keep their steady values from (4.2) (ε replaces Ω). Thus P,N are 'adiabatically eliminated' from the theory on the ground that they reach equilibrium instantaneously on the rapid time scale $T_1 \approx \frac{1}{2}T_2$: the time scale $(C\kappa)^{-1} >> T_1$.

Equation (4.6) shows there is a 'threshold' at

$$G \equiv 1 - 4C\Lambda[1+\varepsilon^2 T_1 T_2]^{-1} = 0 \tag{4.8}$$

where $\Lambda = 1/4C$. Below threshold $G > 0$, and $\partial\varepsilon/\partial t < 0$ so any field decays to zero. If the system is pumped incoherently but strongly so that $\Lambda > 0$ and $G < 0$: then $\partial\varepsilon/\partial t > 0$ and the system settles on a steady non-zero value for ε given by

$$\varepsilon^2 = (4C\Lambda - 1)(T_1 T_2)^{-1} \;, \tag{4.9}$$

where $4C\Lambda > 1$. There are now two real roots for ε, namely $\pm|\varepsilon| \neq 0$. Since ε is a coherent field, coherence has emerged from incoherent pumping. This is the action of the (single mode) laser incoherently pumped above the threshold at $\Lambda = 1/4C$.

The threshold at $\Lambda = 1/4C$ defines a simple 'pitch fork type' bifurcation in which, as the control parameter Λ increases, the three roots of (4.4) move from $\varepsilon = 0$ stable and two non-zero unstable roots to two non-zero stable roots $\varepsilon = \pm|\varepsilon| \neq 0$ and $\varepsilon = 0$ unstable. Evidently there is a 'phase transition' at $4C\Lambda = 1$: however it is a "dynamical" that is "a far from equilibrium" phase transition. Nevertheless it has many of the features of a conventional second order phase transition[19,20].

To see this, note first that (4.6) is expressible in terms of a potential function (free energy) Φ by

$$\begin{aligned} \partial\varepsilon/\partial t &= -\partial\Phi/\partial\varepsilon \quad ; \\ \Phi &\equiv \tfrac{1}{2}c\kappa\varepsilon^2 - 2c\kappa C\Lambda(T_1T_2)^{-1}\ln[1+\varepsilon^2T_1T_2] \\ &\approx +\tfrac{1}{2}c\,\kappa(1-4C\Lambda)\varepsilon^2 + c\kappa C\Lambda(T_1T_2)\varepsilon^4 \quad . \end{aligned} \tag{4.10}$$

We shall extend (4.6) by including an *external* probe field y so that

$$\partial\varepsilon/\partial t = c\kappa\{y - \varepsilon[1 - 4C\Lambda(1+\varepsilon^2T_1T_2)^{-1}]\} \quad . \tag{4.11}$$

In this view $\varepsilon - y$, not ε, is the total field damped inside the cavity: plainly since y is external it is not damped. The coupling of external fields can be quite subtle since their modes are different from those of the cavity. Still (4.11) correctly introduces a qualitatively new feature — coherent pumping by y.

We now have three "thermodynamical" parameters $y, x \equiv \varepsilon$, and $C\Lambda$. Reference to (4.10) shows that $\varepsilon(\equiv x)$ is the 'order parameter'. We identify

$$y \longleftrightarrow (p-p_c) \quad ; \quad \varepsilon \equiv x \longleftrightarrow (V-V_c) \quad ; \quad 4C\Lambda \longleftrightarrow T \quad . \tag{4.12}$$

The quantities p_c, V_c are critical 'pressure' and 'volume'; $T_c = 1$ so that threshold is $T = T_c$. The laser operates on the line $y = 0$ $(p = P_c)$. Here

$$\Phi \approx \tfrac{1}{2}c\kappa(T_c - T)\varepsilon^2 + \tfrac{1}{4}c\kappa(T_1 T_2)T_c\varepsilon^4 \tag{4.13}$$

so that $\partial\Phi/\partial\varepsilon = 0$ means

$$-(T_c - T)\varepsilon = (T_1 T_2)T_c\varepsilon^3 \tag{4.14}$$

and $\partial^2\Phi/\partial\varepsilon^2 = (T_c - T) + 3(T_1T_2)T_c\varepsilon^2$. Thus $\varepsilon = 0$ is the stable root for $T < T_c$. But for $T > T_c$

$$\pm|\varepsilon| = \pm(T_1T_2)^{-1/2}[(T/T_c) - 1]^{1/2} \tag{4.15}$$

are stable non-zero roots.

The behaviour $[(T/T_c) - 1]^{1/2}$ near $T = T_c$ defines a 'critical exponent' $\beta = 1/2$ [20]. Moreover, the equation of state is $(p = -(\partial\Phi/\partial V)_T \longleftrightarrow y = -(\partial\Phi/\partial\varepsilon)_T)$

$$y = c\kappa\{(T_c - T) + T_1T_2T_c\varepsilon^2\}\varepsilon \tag{4.16}$$

the compressibility $\chi = -V^{-1}(\partial V/\partial_p)_T$ is given from

$$\begin{aligned}(\partial y/\partial\varepsilon)_T &= -c\kappa\{(T_c - T) + 3T_1T_2T\varepsilon^2\} && ; \\ &= -c\kappa(T_c - T) \qquad , \ T < Tc && ; \\ &\approx -2c\kappa(T - T_c) \qquad , \ T > T_c && ;\end{aligned} \tag{4.17}$$

so $|(\partial\varepsilon/\partial y)_T|$ jumps from $[c\kappa(T_c - T)]^{-1}$ below threshold to $\tfrac{1}{2}[c(T - T_c)]^{-1}$ above. The exponent -1 defines a critical exponent $\gamma = +1$ [20]. There is an entropy

$$\begin{aligned}S = -(\partial\Phi/\partial T)_V &= S_o - \tfrac{1}{2}c\kappa\varepsilon^2 \\ &\approx S_o - \tfrac{1}{2}c\kappa(T_1T_2)^{-1}[(T/T_c) - 1]\end{aligned} \tag{4.18}$$

above threshold and $S = S_o$ at threshold. $S = S_o$ just below threshold too so S is continuous. But the 'specific heat' $C_v = -T(\partial S/\partial T)_v$ is $\frac{1}{2}c\kappa(T_1T_2)^{-1}$ for $T > T_c$ and vanishes for $T < T_c$. There is therefore a jump that defines the critical exponent $\alpha = 0$[20]. The three critical exponents are therefore $\alpha = 0$, $\beta = \frac{1}{2}$, $\gamma = 1$, and satisfy[20]

$$\alpha + 2\beta + \gamma = 2 \quad . \tag{4.19}$$

These features characterize the so called 'mean field' theories of equilibrium phase transitions[20]. (This 'mean field' has little to do with the spatially averaged mean field theories of optics as in Ref. 18 referred to above).

This positive pumping ($\Lambda > 0$) theory extends easily to several related models. For example a "saturable absorber". This is simply a negative pumping term and is derived from the steady value of N in (4.2) with pumping $\Lambda \to \Lambda'$ and $-1 \le \Lambda' < 0$. Set $M' = 1 - \Lambda'$ so $0 < M' \le 2$. There will be time constants T_1' and T_2' (say) so (4.11) becomes

$$\begin{aligned} \partial\varepsilon/\partial t = \; & c\kappa\{y - \varepsilon[1 - 4C\Lambda(1+\varepsilon^2T_1T_2)^{-1} \\ & - 4C'(1-M')(1+\varepsilon^2T_1'T_2')^{-1}] \quad . \end{aligned} \tag{4.20}$$

Note that two sets of dipoles and inversions are now adiabatically eliminated.

For $y \neq 0$ we have the laser with saturable absorber and injected signal. If $y = 0$ the system is capable of displaying either a first or a second order phase transition as well as 'tricritical phenomena'. In the last a line for first order transition points meets a line of second order transition points at a point specified by $C = C_o$, $C' = C_o'$. This case is noted by Argawal[21]. If $y \neq 0$ 'tristability' is possible: there can be three competing minima of Φ as shown in Fig. 1 (taken

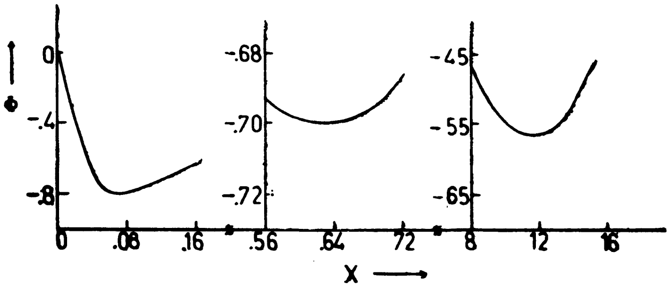


Fig. 1. An example of tri-stability. There are three competing minima of Φ (taken from Ref. 21).

from Ref. 21: $\varepsilon \equiv x$ in Fig. 1). A feature in this case $y \neq 0$ is that symmetry (for $\pm|\varepsilon|$) is now broken by y and $-\varepsilon$ is in phase with y. Agarwal[21] also shows both bicritical and tetracritical behaviour are possible for the two-mode laser coupled in the form

$$\begin{aligned} \dot{\varepsilon}_1 &= (a_1 - |\varepsilon_1|^2 - |\varepsilon_2|^2)\varepsilon_1 \\ \dot{\varepsilon}_2 &= (a_2 - |\varepsilon_1|^2 - |\varepsilon_2|^2)\varepsilon_2 \quad . \end{aligned} \tag{4.21}$$

In tetracritical behaviour four lines of second order phase transitions meet at a point, the tetracritical point (See Fig. 2). In bicritical behaviour one of the four phases disappears and is replaced by a first order transition. There is a co-existence region between two of the remaining phases, and one line of first order transitions and two of second order transitions meet at a point, the bicritical point (see Fig. 3).

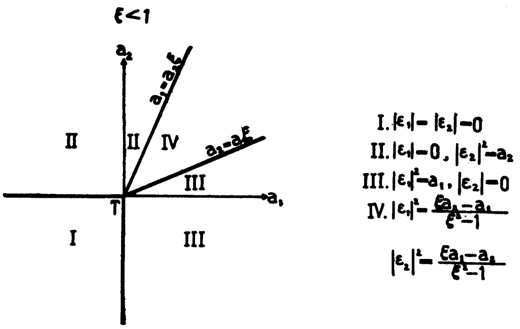

ξ<1
a_2
$a_1=a_2\xi$
$a_2=a_1\xi$
II
II
IV
III
T
a_1
I
III
I. $|\varepsilon_1| = |\varepsilon_2| = 0$
II. $|\varepsilon_1| = 0,\ |\varepsilon_2|^2 = a_2$
III. $|\varepsilon_1|^2 = a_1,\ |\varepsilon_2| = 0$
IV. $|\varepsilon_1|^2 = \frac{\xi a_2 - a_1}{\xi^2 - 1}$
$|\varepsilon_2|^2 = \frac{\xi a_1 - a_2}{\xi^2 - 1}$

Fig. 2. Tetracritical behaviour shown by the two-mode laser equations (4.21). Four lines of second order phase transitions meet at a tetracritical point (taken from Ref. 21). The symbols refer to eqn. (4.21).

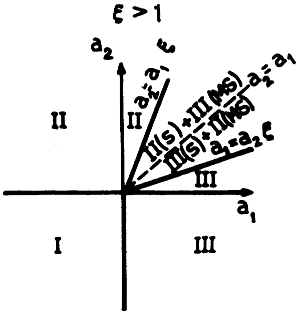


Fig. 3. Bicritical behaviour shown by the two-mode laser equations (4.21). One of the four phases in Fig. 2 has disappeared and is replaced by a first order transition with a co-existence region. One line of first order and two lines of second order now meet at a point - the bicritical point: S and MS denote 'stability' and 'metastability'.

These phase transitions all arise as simple bifurcations as control parameters are changed. It is now known that a spectacular *sequence* of such bifurcations may be associated with the envelope M-B equation which may eventually evolve to 'chaotic' (or turbulent) motion.

Consider the damped laser equations we have obtained with $y = 0$ and P,N, no longer adiabatically eliminated. There is no spatial variation by assumption. So we have from (4.5) and (4.1) at exact resonance that

$$\frac{d\varepsilon}{dt} = c\alpha P - c\kappa\varepsilon$$

$$\frac{dP}{dt} = \varepsilon N - T_2^{-1} P$$

$$\frac{dN}{dt} = T_1^{-1} (\Lambda - N) - \varepsilon P \tag{4.22a}$$

where $\Lambda > 0$. As first noted by Haken[22], this reduces to a standard form of the Lorenz model[23,24,25].

$$\dot{X} = \sigma Y - \sigma X$$

$$\dot{Y} = XZ - Y$$

$$\dot{Z} = b(r-Z) - XY \tag{4.22b}$$

if $\sigma = c\kappa T_2$, $\tau = tT_2^{-1}$, $X = T_2\varepsilon$, $Y = T_2\alpha\kappa^{-1}P$, $Z = T_2\alpha\kappa^{-1}N$, $r = T_2\alpha\kappa^{-1}\Lambda = 4C\Lambda$ and $b = T_2T_1^{-1}$; $\dot{X} = dX/d\tau$ etc. (The more usual form of the Lorenz equations uses $r - Z \rightarrow + Z$, ref.[24].

Lorenz[23] first introduced his equations in connection with the Rayleigh-Bénard convective instability: a liquid between two horizontal plates heated from below develops two dimensional structure as the parameter r increases. For $r < 1$, $X = Y = 0$, $Z = r$ is the stable steady solution. For $r = 1$ ($4C\Lambda = 1$) there is a simple pitch fork bifurcation to the two stable steady values

$$(\pm\sqrt{b(r-1)},\ \pm\sqrt{b(r-1)}, 1) \tag{4.23}$$

and (0, 0, 1) is now unstable. In the Bénard instability the structure consists of 'rolls' normal to the two dimensional spatial plane of right handed or left handed convective flows[24,25]. These flows are stable for $r < r_2$. For $r > r_2$ there are no stable steady state solutions; and for much of $r > r_2$ computer experiments show no stable periodic solutions either. They do not settle down to steady or periodic flow, and show no obvious predictable long range pattern. Small changes in initial data result in substantially different solutions. This is the 'strange attractor' associated with the Lorenz equations: examples of such attractors are shown below in Figs. 8(b), (c) whilst Fig. 11(b) below shows the strange attractor associated with the output from an optical cavity. Fig. 12(b) below shows the spectrum (time Fourier transform) of this strange attractor and 13(b) its evolution. Strange attractors are characterised by their capacity to attract, i.e. flows move toward them. They are not 'fixed points' (that is simply stable steady states) nor are they simple attractive surfaces since the motions display a complicated braided internal structure (Figs. 9(e), (f)) which shows the same features on successively reducing scales (successive magnifications): they are certainly 'strange' (the terminology following Ruelle and Takens[26]).

In a small range of $r < r$, i.e. $r_1 < r < r_2$ the X,Y,Z space divides into two regions. For initial data in one, the system moves to one of two steady states, while for initial data in the other, the motion remains in that region oscillating irregularly apparently forever. The commonly used values of σ and b are $\sigma = 10$, $b = 8/3$ from which $r_1 \approx 24.06$ and $r_2 \approx 24.74$. For $r < r_1$ there are still exceptional trajectories which oscillate irregularly though most initial data eventually settle down.

At $r = r_2 \approx 24.74$ the two stable fixed points (4.23) move to instability by a Hopf bifurcation. Typically in a Hopf bifurcation the system moves to stable *oscillation* as the flows move off the fixed point to a limit cycle. In practice the motion is turbulent on the strange attractor. For $24.06 < r < 24.74$ the two fixed points are stable but simply largely inaccessible. Remarkably enough a crucial bifurcation of collosal proportions (a 'homoclinic explosion'[27]) occurs at $r = 13.926$ where enormous numbers of unstable periodic motions are created. Since they are unstable they cannot easily be seen numerically, but they allow the bifurcation sequence for increasing r to involve bifurcation to a *decreasing* number of states; that is unstable periodic motions can move to mutually annihilate similar stable ones. It is not known to the author what characterises the fundamental frequencies of the periodic orbits. Each frequency (each period $2^n T$) may bifurcate by period halving down to that basic period T. This is the opposite of period doubling. Period doublings associated with a possible accumulation point and a move to 'chaos' on a strange attractor are associated with the name of Feigenbaum who studied mappings[28]. An example of period doubling is shown in Fig. 8(a) below. Period triplings etc. are also possible. What determines what is an open question at present. Certainly the dynamics of the Lorenz system (4.22b) involves an extremely rich behaviour. The reader is referred to Sparrow[27] for the most comprehensive analysis to date.

It will be plain from (4.22a) for the laser that the corresponding equations for the laser with saturable absorber will be

$$\frac{d\varepsilon}{dt} = c(\alpha P_1 + \beta P_2) - c\kappa\varepsilon \qquad ;$$

$$\frac{dP_1}{dt} = \varepsilon N_1 - T_2^{-1} P_1 \quad ; \quad \frac{dN_1}{dt} = T_1^{-1}(\Lambda - N_1) - \varepsilon P_1 \quad ;$$

$$\frac{dP_2}{dt} = \varepsilon N_2 - T_2'^{-1} P_2 \quad ; \quad \frac{dN_2}{dt} = T_1'^{-1}(\Lambda' - N_2) - \varepsilon P_2 \quad . \qquad (4.24)$$

Under the mapping $\sigma = c\kappa T_2$, $\tau = tT_2^{-1}$, $\varpi = T_2 T_1^{-1}$, $\eta = T_2 T_2'^{-1}$, $a = \sqrt{T_1 T_2}\varepsilon$ $b = \sqrt{T_1 T_2} r_1^{-1} \alpha\kappa^{-1} P_1$, $c = 1 - T_2 \alpha\kappa^{-1} r_1^{-1} N_1$, $d = \sqrt{T_1 T_2} r_2^{-1} \beta\kappa^{-1} P_2$, $e = 1 - T_2 \beta\kappa^{-1} r_2^{-1} N_2$, $r_1 = T_2 \beta\kappa^{-1} \Lambda'$ these equations become

$$\begin{aligned} \dot{a} &= \sigma(-a + r_1 b + r_2 d) \\ \dot{b} &= -b + (1-c)a \quad ; \quad \dot{d} = -\eta d + (1-e)a \\ \dot{c} &= \varpi(-c + ab) \quad ; \quad \dot{e} = \varpi(-\lambda\eta e + \mu a d) \end{aligned} \qquad (4.25)$$

with $\dot{a} \equiv da/d\tau$, etc. The undefined parameters are $\lambda = T_1 T_2^{-1} T_2' T_1'^{-1}$ and $\mu = T_1^2 T_2^{-1} T_2'^{-1}$. For radiation damping $T_1 T_2^{-1} = T_1' T_2'^{-1}$ so $\lambda = 1$ whilst $\mu = \eta \varpi^{-2}$ so is not free. We set $r_1 = r_T > 0$ and $r_2 = -r_S < 0$ (laser with absorber), relabel η as a new parameter τ and interpret $\dot{a}$ as da/dt, etc. for some scaled time t which is not the new parameter τ. Providing we can set $\mu = 1$, these equations become

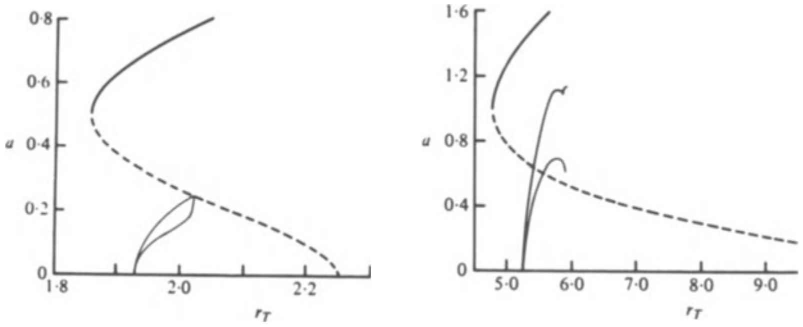
$$\begin{aligned} \dot{a} &= \sigma(-a + r_T b - r_S d) \\ \dot{b} &= -b + (1-c)a \quad ; \quad \dot{d} = -\tau d + (1-e)a \\ \dot{c} &= \varpi(-c + ab) \quad ; \quad \dot{e} = \varpi(-\tau e + ad) \ . \end{aligned} \qquad (4.26)$$

In this form, but without the constraint $\mu = 1$, the equations are precisely the equations of thermosolutal convection studied by Da Costa, Knobloch and Weiss[29]. In thermosolutal convection Rayleigh-Bernard convective flow is opposed by a solute moving under gravity. The parameters r_T and r_S are thermal and solutal Rayleigh numbers. If $r_S = 0$, the equations decouple to another form of the Lorenz equations (see 4.22b). Da Costa et al. studied the case, $\sigma = 10$, $\tau = 0.4$ and $\varpi = 8/3$. This corresponds to the case usually studied for Lorenz equations with $\sigma = 10$, $b = 8/3$(Lorenz parameter). At the time of this writing I have not pursued the case of interest in optics equations (4.25) with $\mu = \eta\varpi^{-2}$. We shall therefore append here to the results reported by DaCosta et al. based on (4.26) with their choice $\sigma = 10$, $\tau = 0.4$ and $\varpi = 8/3$. I hope to report on (4.25) elsewhere.

For arbitrary σ, τ, > 0 and fixed $r_S > 0$ (in 4.26) there are a number of thermal Rayleigh numbers r_T of interest. Two important values are $r_T^{(0)}$ and $r_T^{(e)}$ We shall distinguish cases $r_T^{(0)} > r_T^{(e)}$ and $r_T^{(0)} < r_T^{(e)}$ Henceforth we systematically use the notation of[29]. If $r_T^{(0)} > r_T^{(e)}$ there is a simple pitch fork bifurcation of laser phase transition type from a stable fixed point $a = 0$ to two finite amplitude fixed points $a = \pm|a| \neq 0$ ($a = 0$ is unstable). The two fixed points $\pm|a|$ are stable if, near $r_T^{(e)}$,

$$r_T = r_T^{(e)} + a^2 r_2^{(e)} + O(a^2) \ . \tag{4.27}$$

and $r_2^{(e)} > 0$ all three fixed points are unstable in the sense the motion still moves from $a = 0$, and two fixed points $\pm |a|$ are unstable, if $r_2^{(e)} < 0$.



(a) (b)

Fig. 4. Plot of a against thermal Rayleigh number r_T for $r_S = 0.4$(a) and $r_S = 4$(b). The solid heavy line indicates the stable steady branch (branch of stable fixed points). The broken heavy line is the unstable steady branch (branch of unstable fixed points). The solid thin lines denote the maximum and root mean square values of a on the oscillatory branch.

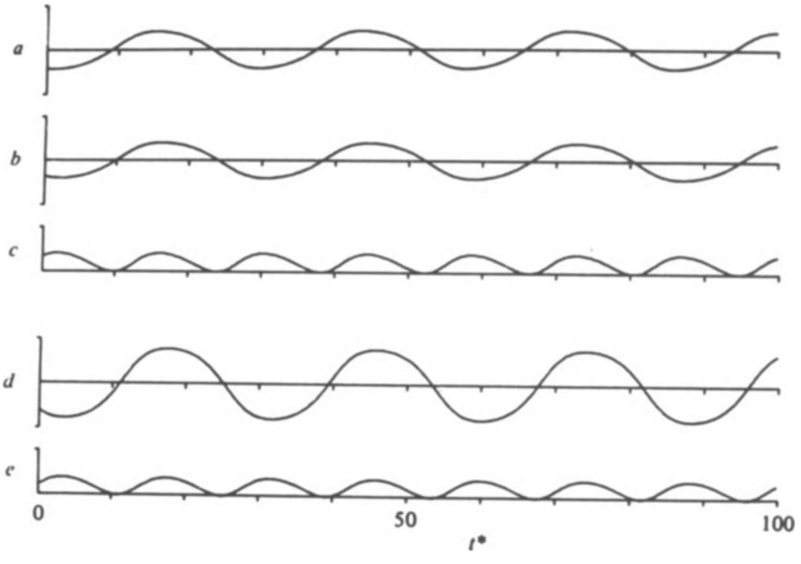


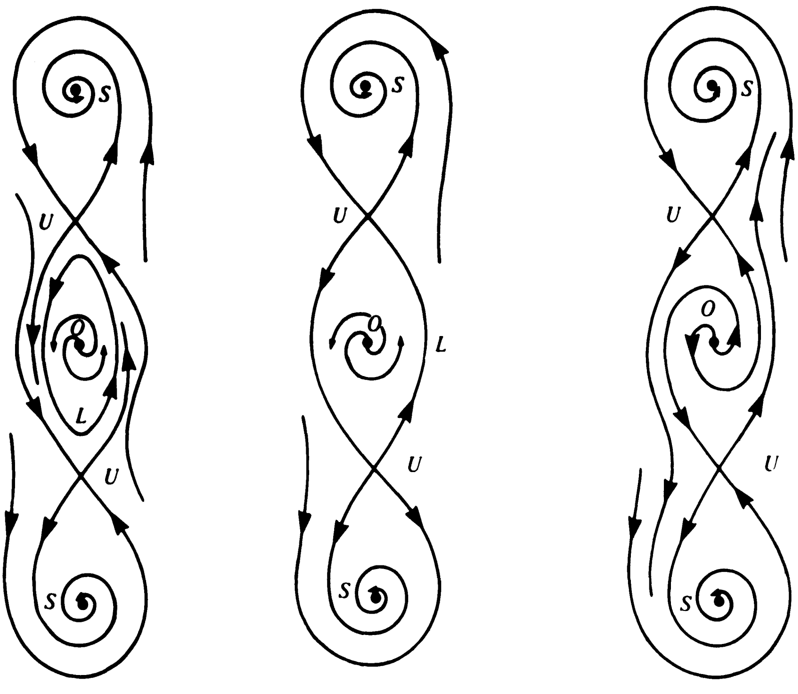
Fig. 5. Typical sets of oscillatory solutions for $r_S = 0.4$ with $r_T = 2.0$. The modes c,e (which are essentially inversions N_1, N_2, see text) have half the period of the other modes. The half-lengths of the ordinate axes are 0.5 for a,b,d,e and 0.1 for c.

If $r_T^{(0)} < r_T^{(e)}$ there is a Hopf bifurcation at $r_T = r_T^{(0)}$ and the system moves off the unstable fixed point $a = 0$ into oscillation. The case of greatest interest combines $r_T^{(0)} < r_T^{(e)}$ and $r_2^{(e)} < 0$. This situation is shown in Fig. 4(a) (for small $r_S = 0.4$) where there is a stable upper branch and a turning point to instability on a lower (dotted branch) at $r_T^{(min)} \approx 1.86$ (only one set of fixed points $a = +|a| \neq 0$ is shown). The oscillatory branch bifurcates at $r_T^{(0)}$ and terminates on the unstable lower branch at $r_T = r_T^{(c)}$. At this point the system jumps to the upper stable branch and there is hysteresis down to $r_T^{(min)}$ if r_T is reduced. At $r_T^{(min)}$ a jumps back to the stable fixed point at $a = 0$. The Fig. 5 shows the type of oscillations which occur for small $r_S = 0.4$. Quantities c,e are roughly rectified forms of a, so that if a has period T, c and e have period $\frac{1}{2}T$. In terms of the inversions N_1, N_2 these oscillate at $\frac{1}{2}T$ compared with the period T of the field ε. The point is that equations (4.25) (or (4.24)) are invariant under $a \to -a$, $b \to -b$, $d \to -d$, $c \to c$, $e \to e$, and if this property applies to the solutions the result follows. We shall see shortly that this property does not necessarily apply to the solutions which though symmetric for small values of r_S, become asymmetric (Fig. 7b below) as r_S is increased.

The symmetric oscillatory branch of Fig. 4(a) is exemplified by the phase plane plot $\dot{a}$ against a shown in Fig. 6. The two stable fixed points marked S are asymmetrically placed about the unstable fixed point a=0 marked 0. The oscillation is described by the limit cycle marked L and the two symmetrically placed unstable fixed points are marked U. For $r_T^{(0)} < r_T^{(c)}$ (Fig.6(a)) the limit cycle is in evidence and the motion has moved by Hopf bifurcation off the now unstable fixed point at $a = 0$. For $r_T = r_T^{(c)}$ (Fig. 6(b)) the limit cycle has moved to the separatrices which join the unstable fixed points U. In these circumstances the period T of oscillation is necessarily infinite. For $r_T^{(c)} < r_T < r_T^{(e)}$ the limit cycle (and all oscillation) has disappeared and the motion spirals in to one or other of the stable fixed points (Fig. 6(c)).

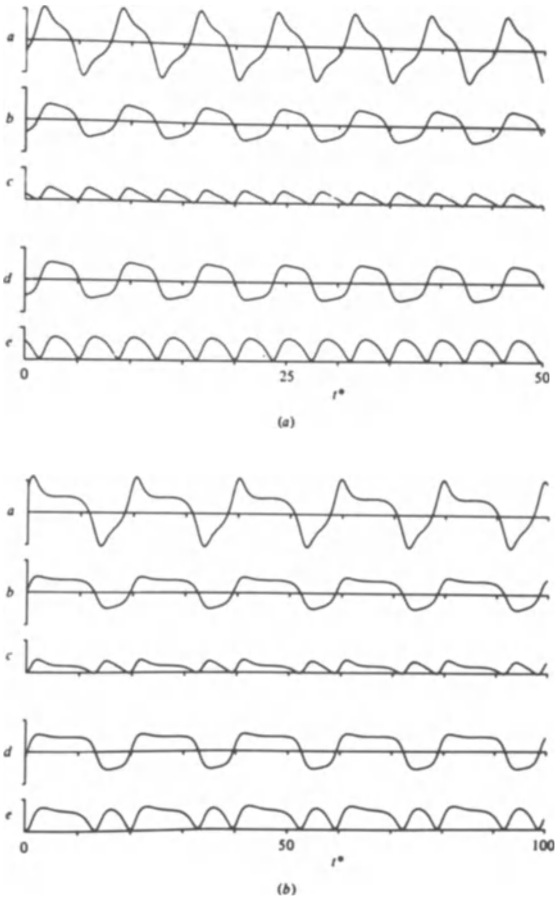
The main effect from increasing r_S is that the amplitude a increases, typically beyond the value on the unstable steady branch (branch of unstable fixed points). Fig 4(b) shows the situation for $r_S = 4.0$. As r_T increases a reaches a maximum at $r_T \approx 5.8$ and at $r_T = 5.855$ the oscillations bifurcate to asymmetrical solutions where a(t) spends more time on one side of $a = 0$ than the other. The period T is not changed. But since c,e are rectified forms of a their periods must double from $\frac{1}{2}T$ to T. Symmetric oscillations are shown in Fig. 7(a) for $r_T = 5.853$. At $r_T = 5.8722$ this has become the asymmetric oscillation shown in Fig. 7(b). At $r_T = 5.874$ this oscillatory solution loses stability and spirals into the fixed point.

At higher $r_S (r_S = 6)$ period doubling and chaotic behaviour occur. There is a bifurcation to asymmetry at $r_T \approx 7.785$ and a bifurcation to double period at $r_T \approx 7.793$. The corresponding limit cycle is shown in Fig. 8(a) and, for the region marked, in detail in Fig. 9(a). The



(a) (b) (c)

Fig. 6. Sketches of the solution trajectories projected onto the $\dot{a},a$ plane for $r_S = 0.5$ and (a) $r_T^{(0)} < r_T < r_T(c) = 2.02195$; (b) $r_T = r_T(c)$; (c) $r_T(c) < r_T < r_T(e)$. O denotes the static solution, U the unstable steady solutions, and L the limit cycle. The notations are explained in the text and in ref. 24, from which this Figure is taken. The parameter r_S is too small for chaos which is seen in Figs. 4 and 5 for $r_S = 6$.



(a)

(b)

Fig. 7. Oscillatory solutions of equations (4.26) for $r_S = 4$ and (a) $r_T = 5.853$. The oscillations in a,b,c,d,e are all symmetrical, 1-c and 1-e act like the inversions N_1, N_2 of equations (4.24). For (b) $r_T = 5.8722$ the oscillations are not symmetrical and there is apparent period doubling in c and e. The half-lengths of the ordinates are 1.0 for a,b,d,e and 0.1 for c.

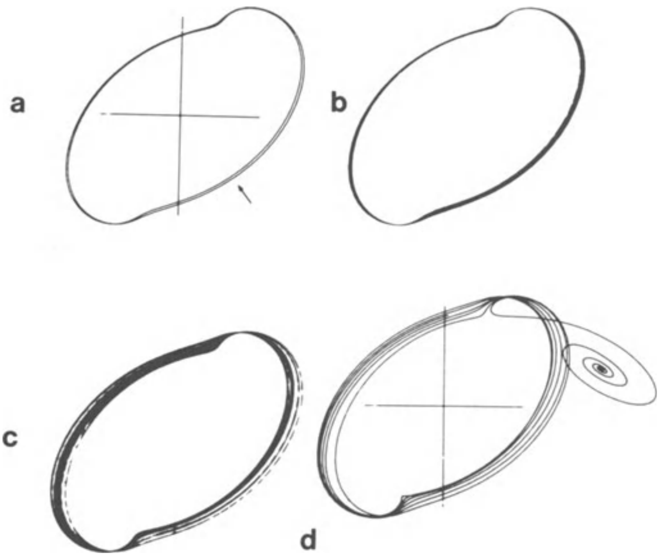


Fig. 8. Limit cycles for equations (4.24) projected onto the a-d plane (a) Period doubling for $r_T = 7.7948$; (b) Chaotic solutions for $r_T = 7.796$; (c) Chaotic solution for $r_T = 7.898$; (d) $r_T = 7.8$; the attractor loses its stability and the trajectory spirals into a fixed point. The axes have half lengths 1.0 and 0.5. Parameters in (4.24) are $J = 10$, $\tau = 0.4$, $= 8/3$ and $r_S = 6$ (taken from Ref. 29).

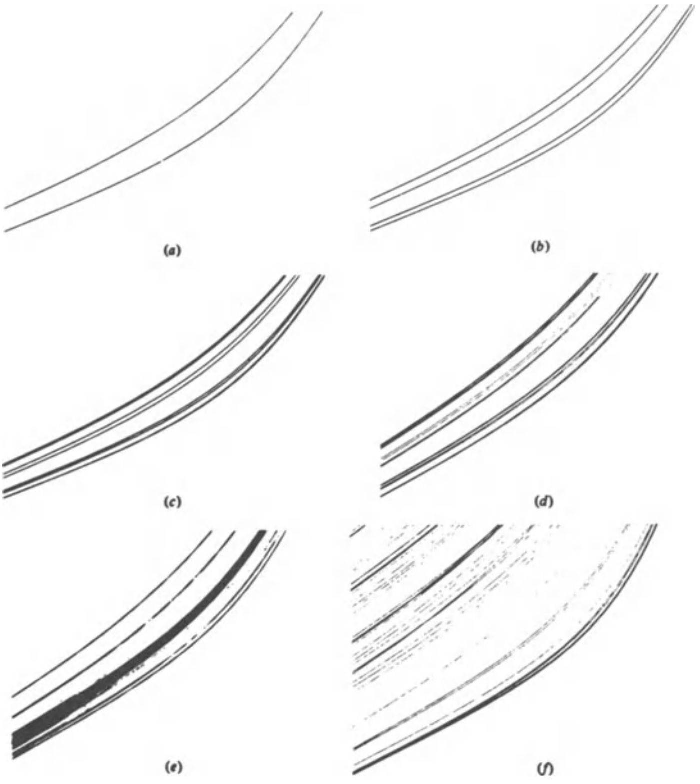


Fig. 9. Details of the limit cycles in Fig. 8 for the region indicated in Fig. 8(a) $r_T = 7.7948$; (b) $r_T = 7.952$ (period quadrupling); (c) $r_T = 7.79545$ (octupling); (d) $r_T = 7.79548$ (16-fold); (e) $r_T = 7.796$ (chaotic); (f) $r_T = 7.798$ (chaotic). The scale for f differs from that for (a) to (e) (taken from Ref. 29).

two segments are traversed alternately. At $r_T \approx 7.7949$ period doubling leads to the four segments shown in Fig. 9(b). Further period doublings 2^nT occur at $r_T \approx 7.7953$ $(n = 3)$, $r_T \approx 7.79546$ $(n = 4)$ and $r_T = 7.9550$ $(n = 5)$. The solutions for $n = 3, 4$ are shown in Figs. 9(c), (d) (the apparently discrepant numbers given in the caption to Fig. 9 are taken from Ref. 29 as of course is the whole of this discussion). For $r_T > 7.7955$ there are no periodic solutions and the motion is on a strange attractor. Figs. 8(b), (c) show such motions for $r_T = 7.796$ and 7.798. The former shows motion in the neighbourhood of one (unstable) asymmetric oscillation. The latter moves between such a one and its mirror image. In both cases the motions have a complicated braided structure as shown in Figs. 9(e), (f). These braidings are expected to present themselves on ever diminishing scales. For $r_T > 7.8$ these attractors lose their stability and the motion spirals into a fixed point as shown in Fig. 8(d).

Period doublings at values $r_T^{(n)}$, with an accumulation point at $\lim_{n\to\infty} r_T^{(n)}$ with chaotic motion beyond that limit point seems to have universal character. Feigenbaum[28] showed that for a wide class of difference equations the quantity analogous to

$$\delta_{n-1} = \frac{r_T^{(n)} - r_T^{(n-1)}}{r_T^{(n+1)} - r_T^{(n)}} \qquad (4.27)$$

tends to $\lim_{n\to\infty} \delta_n = 4.6992 \cdots$. Weiss and colleagues in Ref. 29 find $\delta_1 = 2.32$, $\delta_2 = 4.32$, $\delta_3 = 4.58$, $\delta_4 = 4.67$. Our demonstration that the equations (4.26) for thermosolutal convection are a particular case of the laser equations (4.25) with saturable absorber, and the equivalence of the laser equations (4.22a) to the Lorenz equations, shows that the damped M-B equations have a universal character which is not peculiar to lasers with positive pumping (or with positive and negative pumping as in (4.25) in their envelope forms (4.5) with (4.1) (in which spatial variation is eliminated).

5. SEMICLASSICAL THEORY OF OPTICAL BISTABILITY

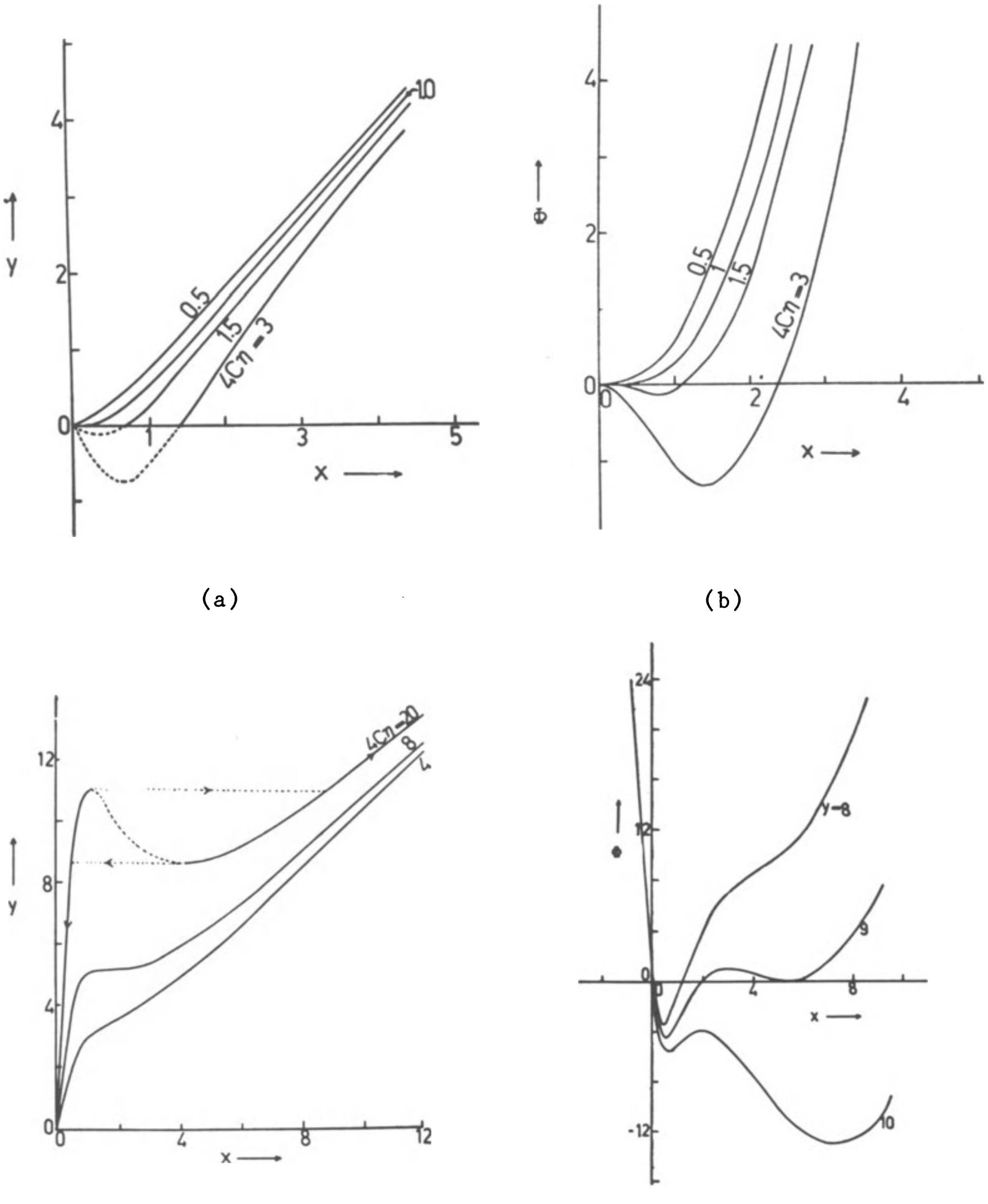
Here we consider (4.11) with $\Lambda < 0$. If the atoms start in their ground states $\Lambda = \underline{-1}$. It is convenient to set $-4C\Lambda = +4C\eta$. We can scale y and $x = \varepsilon\sqrt{T_1 T_2}$ to get the steady state in the form

$$y = x + 4C\eta\, x / [1 + x^2] \quad , \qquad (5.1)$$

and this is plotted for different values of $C\eta$ in Fig. 10(c). For $4C\eta < 8$ there is only one stable state; for $4C\eta = 8$ there is a point of inflexion where $dy/dx = 0$ and $d^2y/dx^2 = 0$. Beyond this there are two stable states (= minima of Φ) as shown in Fig. 10(d) but the system is monostable for still larger $C\eta$. For $4C\eta = 10$ the Fig. 10(c) shows two stable branches for x against y and an unstable branch connecting them is shown dotted. The dotted line parallel to y = constant shows the jumps, with hysteresis, actually followed as the coherent jumping parameter y is increased and then decreased. This equation of state x against y looks like a V,p plot for a *first* order phase transition. But in the neighbourhood of the critical point $\partial^2 y/\partial x^2 = 0$, $\partial y/\partial x = 0$, with $2C\eta \equiv T$ the transition is equally well described as a second order phase transition[30].

For comparison the Figs. 10(a) and (b) shows comparable results for the equation of state for the laser derived from (4.11) with $\eta = +\Lambda$. The laser operates on $y = 0$ and Fig. 10(a) shows unstable regions dotted. The field $x \propto \varepsilon$ jumps to finite operating values on $y = 0$ for $4C\eta > 1$ and these values minimise Φ as Fig. 10(b) shows.

Note that all of these phenomena have emerged from the envelope



(a) (b)

(c) (d)

Fig. 10. Comparable results for the laser $\Lambda > 0$ and coherent pumping $y \neq 0$ and $\Lambda < 0$: (a) equation of state for the laser derived from equations (4.11) with $\eta \equiv +\Lambda > 0$. Without an injected signal the laser operates on $y = 0$. The dotted regions are unstable. The field X jumps to finite operating values for $4C\eta > 1$ and these values minimise Φ shown in (b). In (c) y and x from equations (5.1) are plotted for different $C\eta = -C\Lambda$ and $\Lambda < 0$. The Fig. shows how two stable states emerge from one. These stable states are minima of Φ shown in (d). The system is again monstable for still larger $C\eta$ (taken from Ref. 21)

B-M equations (4.1) (4.5) without spatial variations. We are on exact resonance: there is incoherent pumping $\Lambda > 0$ and no coherent pumping $(y = 0)$ for the laser; there is incoherent pumping $-1 \leq \Lambda < 0$ and there is coherent pumping $y \neq 0$ for the bistability. The latter is called "optical bistability" and at resonance it is purely 'absorptive' optical bistability. A closely related theory of 'dispersive' optical bistability appeals to the same equations. The atoms must lie inside a cavity whose excited mode is detuned from the wave number ωc^{-1}. In practice the coherent pumping is imposed by a field entering the cavity from outside and the induced field ε is observed as a transmitted field emerging from the cavity. Thus the cavity acts as a Fabry-Perot interferometer and plays an essential role. We return to cavity action (necessarily all too briefly) in discussion of the quantum theory below.

However, it would be incomplete, in view of the Lorenz model considered in §4, if we left the theory of optical bistability, even in the simplified form given in this section, without noting the possibility of actually observing a strange attractor in the laboratory. A ring cavity has the property of admitting light which can then be either simply transmitted or made to follow a track back to the input point from which it can repeat the process. A very much simplified theory goes as follows: The linearised theory of §2 yields the refractive index $m(\omega)$ and its despersion relation (2.12b). This becomes non-linear by noting that the right side of the driven oscillator (2.9b) involves $-2p\hbar^{-1}\omega_s r_z(\underline{x},t)E(\underline{x},t)$. Linearisation puts $r_z = \pm 1$. We might therefore expect that in the non-linear regime at low enough densities n there is a refractive index of the general form[31]:

$$m - 1 = \frac{2\pi n p^2 \hbar^{-1}(1+(\Delta\omega)^2 T_2{}^2)}{(\Delta\omega + iT_2{}^{-1})(1+\Omega^2 T_1 T_2 + (\Delta\omega)^2 T^2)} \tag{5.2}$$

We have substituted $r_z(\underline{x},t) \to N$ where N is given by (4.2) and guessed the corresponding form of N off-resonance. This form is essentially correct[31]. Evidently, for small enough Ω^2, $m - 1 = m_o - 1 + A\Omega^2$ where m_o is the linearised value. It follows that in the path round a ring cavity of effective track length L there is a phase shift (phase delay) $\phi = \omega c^{-1} m L$ and ϕ is of the form $\phi_o + B\Omega^2$. The field at the entry point to the ring cavity consists of the sum of the input E_o and the field E_N which has traversed the ring on its nth loop. Th field is of the form $E_n e^{i\phi_n}$ since ϕ_n depends on E_n as $\phi_n = \phi_o + BE_n{}^2$. We therefore obtain a different equation of the form

$$E_{n+1} = E_o + E_n e^{i(\phi_o + BE_n{}^2)} \quad . \tag{5.3}$$

This map, due to Ikeda[32] goes into a series of bifurcations forming limit cycles of period $2\nu T$, $\nu = 1,2,\cdots$ (another period doubling sequence: $T \sim m_o L c^{-1}$); and after an accumulation point it moves onto

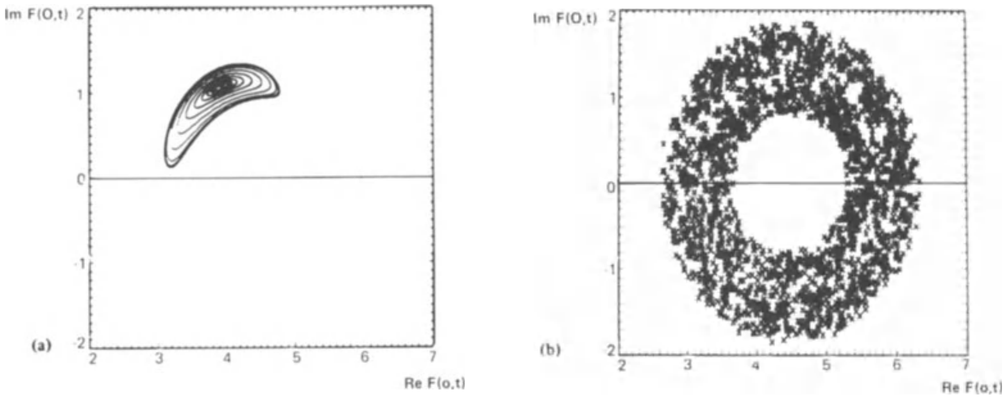


Fig. 11. Strange attractor for the output of a Fabry-Perot cavity (taken from Refs. 33). Successive plots of Im E_n against Re E_n (analoguous to E_n (from (5.3)) fill the region shown in (b). In (a) a Hopf bifurcation moves the system from a fixed point to a limit cycle and oscillation.

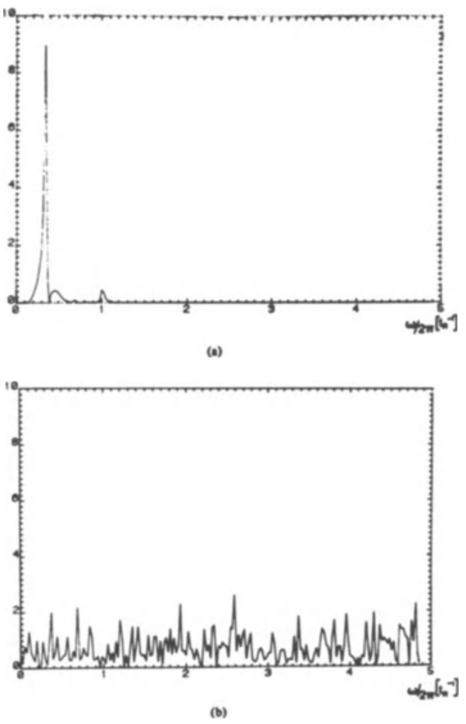


Fig. 12. Spectrum of an oscillatory output from a Fabry-Perot cavity (a) and the comparable spectrum of the strange attractor (b)(taken from Refs. 33).

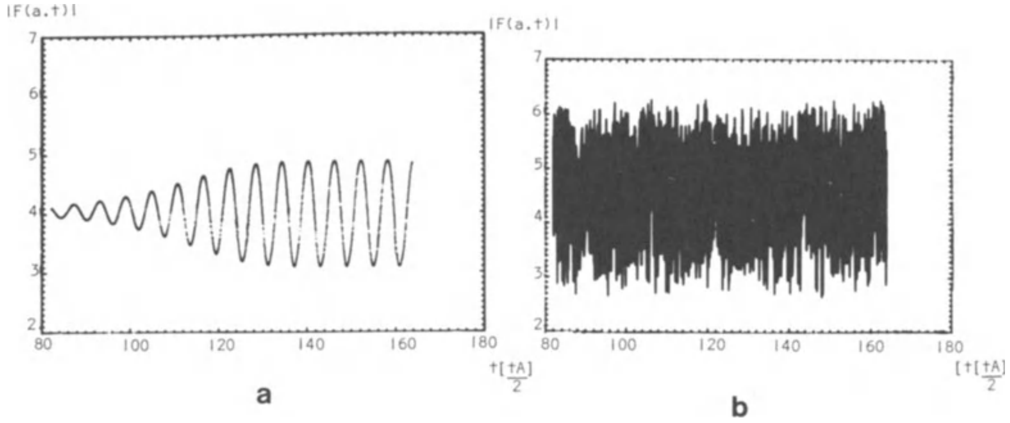


Fig. 13. (a) Time evolution of the oscillatory output shown in Fig. 12(a); (b) time evolution of the strange attractor (taken from Refs. 33).

a strange attractor. The corresponding attractor for a Fabry-Perot cavity is now reported by Firth et al.[33] and their attractor is shown in Fig. 11(b). The Fig. 12(b) shows the power spectrum and Fig. 13(b) the time evolution of what is essentially the output field amplitude. Fig. 13(a) shows one limit cycle. Ikeda's predictions[32] have been confirmed by an experiment using a hybrid optical device[34]

For $n \to \infty$ $E_{n+1} = E_n = E$ (say) and one finds from (5.3) that the input E_0 and the output E are related by a sequence of hysteresis loops rising in position with respect to E as E_0 increases. This means the system is stepwise excited a feature also shown by the Fröhlich oscillator[13], but it is also stepwise de-excited with hysteresis. A slab of dielectric (which acts as a Fabry-Perot interferometer) acts in the same way.

Note that the work reported at this meeting by McLaughlin[35] is also concerned with the ring cavity — namely with 'transverse effects' there. Throughout we have been concerned with propagation in one space dimension z. Effects in the transverse directions, x and y, are also often important — self-focussing for example[5]. In the present case, inside the non-linear medium, (4.11) is replaced by

$$\left(\frac{\partial\varepsilon}{\partial z} - \frac{1}{i}\frac{\partial\varepsilon}{\partial t}\right) - i \ F^{-1}\nabla_T^2\varepsilon = \frac{-2\pi n p^2\hbar^{-1}T_2(1+i\Delta\omega T_2)\varepsilon}{(1+\varepsilon^2T_1T_2+(\Delta\omega)^2T_2{}^2)} \quad . \qquad (5.4)$$

The non-linear form on the right compares with the corresponding expression in (5.2): $\nabla_T^2 \equiv \partial^2/\partial x^2+\partial^2/\partial y^2$ and F is a reciprocal length involving the Fresnel number measuring the transverse diffraction of the transmitted beam (compare Ref. 5). In one transverse direction (x say) this equation is still not integrable in the sense of Ref. 8. It is close to the integrable non-linear Schrödinger equation (refer to §3 just before eqns. (3.5)) and has solitary wave solutions (in x and z) as well as solutions which approximately combine several of these. The transverse intensity along x is then made up of a series of pulses. McLaughlin indicates that these may bifurcate to oscillation with subsequent period doubling. Chaotic motion is therefore expected. Note that the pulsed transverse profile is an example of 'filamentation': normally this involves ultra-intense transmitted powers (certainly > 1 Megawatt cm^{-2}). Here we are concerned with 100 mW cm^{-2} or less. It is interesting to speculate whether such filamentation is important at the still much lower radiation intensities which can be significant to biological systems. Such filamentation seems to underlie the work in Ref. 36.

6. THE QUANTUM THEORY

The simplest quantum model, which is so far the only tractable one, deliberately eliminates spatial effects by placing all the atoms on the same site. This proves to be less drastic than it sounds and close agreement of some of its predictions with the results of recent experiments[37] has been found. Physically the model assumes that a large number of atoms lie within a cubic wavelength λ^3 associated with the wavelength λ of the radiation. Coulomb interactions are ignored. The evolution of the atoms is co-operative and the model was introduced by R.H. Dicke[38] with this in mind. We refer to it as the Dicke model in this paper. In the experiments[37] $\lambda \approx 2\mathrm{mm}$. The model should therefore certainly be applicable for frequencies $10^9 - 10^{11}$ Hz. — namely in the frequency range of the Fröhlich oscillators[13].

Given the single site assumption it is natural to introduce operators for N atoms

$$S_x = \sum_{i=1}^{N} \sigma_x(i) \quad , \quad S_y = \sum_{i=1}^{N} \sigma_y(i) \quad , \quad S_z = \sum_{i=1}^{N} \sigma_z(i) \tag{6.1}$$

and find, through Heisenberg's equations of motion for a Hamiltonian

$$H = \tfrac{1}{2}\hbar\omega_s S_z + H_f - \underset{\sim}{p\hat{u}}\, S_x \cdot \underset{\sim}{e}(t) \quad , \tag{6.2}$$

equations of motion for $\vec{S} = (S_x, S_y, S_z)$. In (6.2) the Hamiltonian of the free atoms is $\frac{1}{2}\hbar\omega_s S_z$: H_f is the Hamiltonian for the free quantised e.m. field; and $\underset{\sim}{e}(t)$ is a quantised electric field operator. Because of the single site assumption we can 'reduce' the problem to one for the evolution of $\vec{S}$ alone. (We use $\vec{S}$ for a spin operator not $\underset{\sim}{S}$).

We follow our previous notation and define

$$S_{\pm} = \tfrac{1}{2}(S_x \pm i\, S_y) \tag{6.3}$$

and use $\frac{1}{2} S_z$ for S_z. The commutation relations are

$$[S_+, S_-] = 2S_z, \quad [S_z, S_\pm] = \pm S_+ \tag{6.4}$$

The simplest equations one can find are *collective* Bloch equations in the form

$$\begin{aligned} \langle \dot{S}_+ \rangle &= 2\gamma_o \langle S_+ S_z \rangle - 2i\Omega \langle S_z \rangle \\ \langle \dot{S}_- \rangle &= 2\gamma_o \langle S_z S_- \rangle + 2i\Omega \langle S_z \rangle \\ \langle \dot{S}_z \rangle &= -2\gamma_o \langle S_+ S_- \rangle - i\Omega[\langle S_+ \rangle - \langle S_- \rangle] \end{aligned} \tag{6.5}$$

which should be compared with (4.1) at exact resonance ($\Delta\omega = 0$). Resonance of the coherent driving field Ω is assumed in (6.5). The difference from (4.1) is that Ω is now an external driving field. There is no envelope Maxwell field equation like (4.3) for Ω and the collective field is already thrown between the atoms via the operator non-linearities S_+S_z, S_zS_- and S_+S_-. The notation is as before that $\Omega = p\hbar^{-1}E_o$ is a Rabi frequency for the coherent driving field amplitude E_o and resonant frequency ω_s. The e.m. field is in a single mode coherent state $|\alpha\rangle$ where $\alpha \propto E_o$ and the mode wave number is $k_s = \omega_s c^{-1}$. The number $\gamma_o = \frac{1}{2}\Gamma_o (=T_2^{-1})$.

The operator non-linearities are new forms of non-linearity more fundamental than those introduced so far. One way to treat them is to write them as products of c-numbers by approximation. Expectation values are taken in (6.5) so we could for example write the expectation value of the operator product $\langle S_+S_z \rangle$ as the product of expectation values

$$\begin{aligned} \langle S_+ S_z \rangle &= \left\langle \sum_{i=1}^{N} \sigma_+^{(i)} \cdot \frac{1}{2} \sum_{i=1}^{N} \sigma_z(i) \right\rangle \\ &\approx \langle S_+ \rangle \langle S_z \rangle \end{aligned} \tag{6.6}$$

and do the same for other nonlinear terms. This sort of procedure is called 'de-correlation' and ignores specific quantum effects which arise as fluctuations about this approximation. We show below, it leads to some features of a phase transition. This particular phase transition emerges in all detail as an exact solution of the quantum theory which we consider below. First, we consider a de-correlation procedure of slightly different character. By the operator algebra

$$\sigma_+^{(i)} \sigma_z^{(i)} = -\tfrac{1}{2}\sigma_+^{(i)} \quad ; \quad \sigma_+^{(i)} \sigma_z^{(j)} = \sigma_z^{(j)} \sigma_+^{(i)} \quad (i \neq j) \ . \tag{6.7}$$

Then we approximate

$$\langle S_+ S_z \rangle = \sum_{i=1}^{N} \langle \sigma_+^{(i)} \sigma_z^{(i)} \rangle + \sum_{i \neq j} \langle \sigma_+^{(i)} \sigma_z^{(j)} \rangle$$

$$\approx -\tfrac{1}{2}\langle S^{+}\rangle + (1+0\ (N^{-1}))\langle S_{+}\rangle\langle S_{z}\rangle \qquad (6.8)$$

Then, if $Y = \sqrt{2}\Omega/i\gamma_{o}$ $X = \{2\Omega - i\gamma_{o}(1+0(N^{-1})\langle S_{y}\rangle\}/(i\tfrac{1}{2}\sqrt{2}\gamma_{o})$, we find the steady state solution

$$\begin{aligned}\langle S_{x}\rangle_{ss} &= \langle S_{y}\rangle_{ss} = \tfrac{1}{2}\sqrt{2N}X\,[1+X^{2}]^{-1}\\ \langle S_{z}\rangle_{ss} &= -\tfrac{1}{2}N[1+X^{2}]^{-1}\end{aligned} \qquad (6.9)$$

and the equation of state

$$Y = X + (N-1)X\,[1+X^{2}]^{-1} \qquad (6.10)$$

which is a form of (5.1)! The external field $Y \propto \Omega$ provides coherent pumping and the system is bistable for large enough (N-1). The term in (6.8) $-\tfrac{1}{2}\langle S_{+}\rangle$ can be identified as normal radiative damping. Thus (6.8) simply places contributions from radiative damping into this term. The remaining interatomic collective interaction is contained in the remaining term (which is a sum $\sum_{i\neq 1}$ of interatomic terms). Thus the physics of this result is wholly reasonable. Nevertheless the particular form of de-correlation epitomised by (6.8) is crucial if simple optical bistability is to be derived. The exact quantum theory gives a different result and a cavity is needed for optical bistability as will become plain shortly.

Still the result has interesting features. Not least of these are the spectra of scattered light shown in Fig. 14. The different cases are associated with the points (A), (B), (C), (D), (E) shown on the equation of state (6.10) $|X|$ against Y. Plot (A) is a broad band Lorentzian which narrows (through 'critical slowing down') near the phase transition as shown by (B), which has a sharp spike. Plot (D) with hysteresis shows essentially the same thing. The three-peaked spectrum on the upper branch at (C) is typical of *incoherent* resonant scattering. On the other hand we can associate the scattering on the lower branch with *coherent* phenomena. The phase transition is from coherence to incoherence (with hysteresis) as coherent pumping Y increases. As usual the branch with negative slope in the equation of state is unstable.

We return now to the de-correlation procedure (6.6) and then go on the exact quantum theory it represents. If we de-correlate all three non-linearities in (6.5) in the style of (6.6) we get

$$\begin{aligned}\langle \dot{S}_{x}\rangle &= 2\gamma_{o}\langle S_{x}\rangle\ \langle S_{z}\rangle\\ \langle \dot{S}_{y}\rangle &= 2\gamma_{o}\langle S_{y}\rangle\ \langle S_{z}\rangle - 2\Omega\ \langle S_{z}\rangle\\ \langle \dot{S}_{z}\rangle &= -2\gamma_{o}[\langle S_{x}\rangle^{2} + \langle S_{y}\rangle^{2}] + 2\Omega\langle S_{y}\rangle\end{aligned} \qquad (6.11)$$

which also may be compared with (4.1). Despite the apparent damping

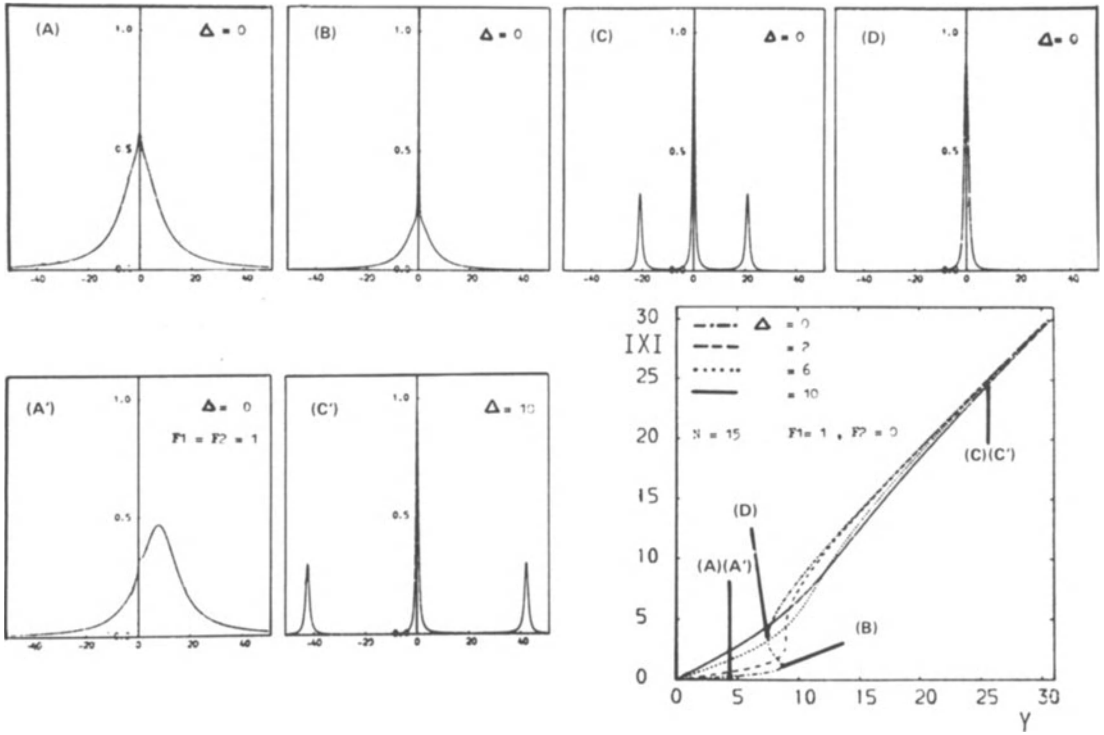


Fig. 14. Spectra of the fluorescent light associated with the de-correlated theory with equation of state (6.10). The different cases (A) - (D) are marked on the plots of this equation of state $|X|$ against y at the bottom right. The coherence of the lower branch and incoherence of the upper branch is epitomised by the different spectra. The numbers Δ, F_1 and F_2 refer to detuning, and certain shifts not described here (taken from S.S. Hassan and R.K. Bullough in Optical Bistability, eds C.M. Bowen, M.Ciftan and H.R. Robl, Plenum Press, New York, 1981).

there is now a constant of the motion $\langle\vec{S}\rangle^2$ and we can define a Bloch sphere radius $N^{-2}\langle\vec{S}\rangle^2 = 1$. There is a solution with $\langle S_x\rangle = 0$. Then if $\langle S_y\rangle = \frac{1}{2}N\sin\sigma$ and $\langle S_z\rangle = -\frac{1}{2}N\cos\sigma$ we need

$$\dot{\sigma} + \gamma_o N\sin\sigma - 2\Omega = 0 \quad . \tag{6.12}$$

This has the steady solution $\dot{\sigma} = 0$ with

$$N^{-1}\langle S_y\rangle \equiv \tfrac{1}{2}\sin\sigma = \Omega\gamma_o{}^{-1}N^{-1} \ (\equiv \tfrac{1}{2}\Theta) \quad . \tag{6.13}$$

Then $N^{-1}\langle S_z\rangle = \pm\frac{1}{2}\sqrt{1-\Theta^2}$ providing $\Theta \leq 1$. For $\Theta \geq 1$, $\langle S_z\rangle = 0$ and the system oscillates in Lotka-Volterra cycles over the whole of the Bloch sphere[39]. One consequence (for $N\to\infty$) is to generate all frequencies $2\Omega, 4\Omega, 6\Omega \cdots$; and these produce an infinite set of side-bands in the spectrum. The 3-peaked spectrum is as before (see (c) in Fig. 14) and the spacing is 2Ω; but there are also side-bands at $\pm 4\Omega$, $\pm 6\Omega$, etc. Evidently there is a bifurcation, with order parameter Θ, such that $N^{-1}\langle S_z\rangle = 0$, $\Theta > 1$ and $N^{-1}\langle S_z\rangle = \pm\frac{1}{2}\sqrt{1-\Theta^2}$, $0 \leq \Theta \leq 1$. The lower branch $-\frac{1}{2}\sqrt{1-\Theta^2}$ must in practice be the stable branch.

This very interesting transition has some features of a second order phase transition, but we have been unable to categorise it completely this way[40]. It does seem to describe what happens in the exact quantum theory quite well, leaving aside effects due to quantum fluctuations, — as we show next. And this seems to mean that the optical bistability found at (6.10) from (6.8) is actually spurious! We show how to get optical bistability from the model shortly.

Instead of the system (6.5), which can be read as a non-linear operator equation, one can always find the equation of motion of the density operator $\hat{\rho}$. This is taken in Schrödinger representation and proves to be

$$\begin{aligned}\gamma_o^{-1}\frac{d\hat{\rho}}{dt} &= g[S_+\hat{\rho} + S_-\hat{\rho} - \hat{\rho}S_+ - \hat{\rho}S_-] \\ &\quad +[2S_-\hat{\rho}S_+ - S_+S_-\hat{\rho} - \hat{\rho}S_+S_-]\end{aligned} \tag{6.14}$$

where $g = -i\Omega\gamma_o{}^{-1}$. Exact steady state solutions of this equation were first given by Puri and Lawande[41]. These have the property

$$\begin{aligned}\hat{\rho}_{ss} &\to (S_-)^2 \ (S_+)^2 \quad , \quad |g| \to 0 \\ &\to I/(2S+1) \quad , \quad |g| \to \infty\end{aligned} \tag{6.15}$$

where $\vec{S}^2 = S_+S_- + S_-S_+ + S_z{}^2$ is a constant of the motion, I is the unit operator, and $S = \frac{1}{2}N$. The steady state solutions are:

$$\langle S_x\rangle = 0, \ N^{-1}\langle S_y\rangle = \tfrac{1}{2}\Theta\left[1 - \frac{N+1}{D}\left(\frac{N\Theta}{2}\right)^{2N}\right]$$

$$N^{-1}\langle S_z\rangle = -\tfrac{1}{2} + \frac{\Theta}{4N}\frac{\partial}{\partial\Theta}(\ln D)$$

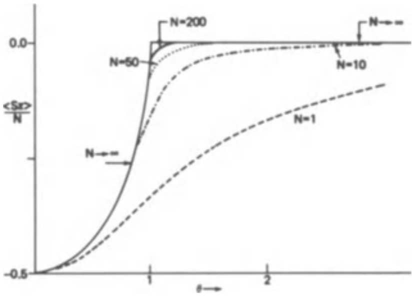


Fig. 15. The solution (6.16) for $m_z \equiv \langle S_z \rangle N^{-1}$ against $\Theta=2|g|N-1$. Note the sharp 'phase transition' for $N \to \infty$.

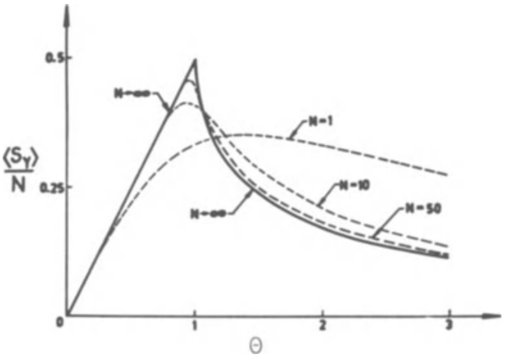


Fig. 16. The solution (6.16) for $\langle S_y \rangle N^{-1}$.

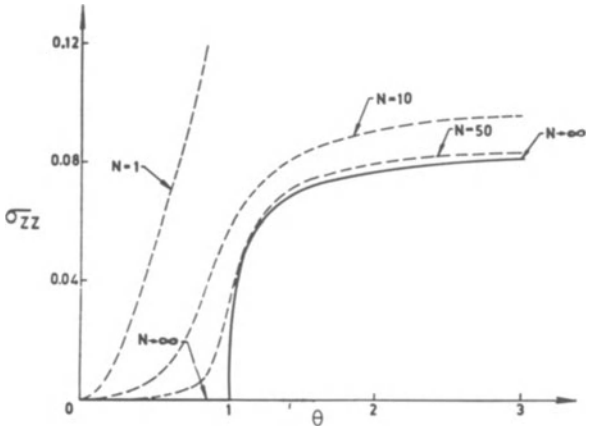


Fig. 17. The quantum fluctuations σ_{zz} for the solution (6.16). There are no quantum fluctuations ($\sigma_{zz}=0$ for $N \to \infty$ and $0<\Theta<1$, and the system is in an atomic coherent state.

where

$$\Theta = 2|g|N^{-1} \quad ; \quad D = \sum_{r=0}^{N} \frac{(N+r+1)!}{(N-r)!}\,\frac{(r!)^2}{(2r+1)!}\,|g|^{2(N-r)} \qquad (6.16)$$

This complicated solution is plotted in Fig. 15 for $m_z = \langle S_z \rangle N^{-1}$ against Θ and in Fig. 16 for $\langle S_y \rangle N^{-1}$ against Θ. For $N \to \infty$ with Θ finite there is a sharp 'phase transition' at $\Theta = 1$.

The nature of *this* phase transition is indicated by the plot of

$$\sigma_{zz} \equiv \{<S_z{}^2> - (<S_z>)^2\} N^{-2} \tag{6.17}$$

which is shown in Fig. 17. Below threshold at $\Theta = 1, \sigma_{zz} = 0$. There are no quantum fluctuations for $N \to \infty$ and one can show that the system is in an atomic coherent state[42]. Above threshold σ_{zz} rises to about 0.08 (for $N \to \infty$) and quantum fluctuations play a rôle in this region. This behavior is missed by the de-correlation (6.6).

One way to find quantal behavior is to look at the scattered radiation. The scattered intensity can be referred to the atoms as sources. A measure of that scattered intensity is $<S_+(\infty)S_-(\infty)> \equiv G^{(1)}(0)$. The argument 0 refers to a time $\tau=0$, but this is a correlation time. The arguments ∞ indicate that the steady state is supposed reached . In principle all the intensity correlations

$$G^{(n)}(0) \equiv <S_+{}^n S_-{}^n> = \mathrm{Tr}(S_+{}^n S_-{}^n \hat{\rho}_{ss}) \tag{6.18}$$

can be measured and so can *correlations* of intensity described by e.g.

$$G^{(2)}(\tau) = \lim_{t \to \infty} <S_+(t)S_+(t+\tau)S_-(t+\tau)S_-(t)> \quad . \tag{6.19}$$

This quantity is usually normalised to $g^{(2)}(\tau) \equiv G^{(2)}(\tau) / \{G^{(1)}(0)\}^2$. Note that $<S_+S_->$ is the intensity because $S_\pm$ are measures of the collective dipole in the system. The theory of quantum measurement shows that all operators must be in the 'normal order' $S_+{}^n$ followed by $S_-{}^n$ in $G^{(n)} \equiv <S_+{}^n S_-{}^n>$.

For the model described here we find for $N = 1$ that

$$\begin{aligned} G^{(1)}(0) &= |g|^2 [1+2|g|^2]^{-1} \\ G^{(2)}(0) &= 0 \quad . \end{aligned} \tag{6.20}$$

We have $g^{(n)}(o) = (S_+^n S_-^n)/\{G^{(1)}(o)\}^n$ and for $N \geq 2$

$$\lim_{|g| \to 0} g^{(n)}(0) = 1, \; n = 1, 2, \cdots, N \tag{6.21a}$$

$$\lim_{|g| \to \infty} g^{(2)}(0) = 1.2 \; (\text{for } N \to \infty) \quad . \tag{6.21b}$$

The quantity $g^{(2)}(0)$ is a key measure of coherence. For a coherent state $g^{(n)}(0) = 1$, for all $n = 1, 2, \cdots$, so (6.21a) indicates a coherent state for $|g| \to 0$ (and $N \to \infty$ for Ω finite). On the other hand black body radiation has $g^{(n)}(0) = n!$, so $g^{(2)}(0) = 2$ and is a measure of incoherence. So we see from (6.21b) that the system is *partially coherent* as $|g| \to \infty$ (strong fields $\Omega \to \infty$ with N). Thus the phase transition is one from coherence to partial coherence as the parameter Θ increases. Notice we actually work at Ω, $N \to \infty$ but Θ finite in order to define a sharp phase transition in 'thermodynamic limit'.

Similar behaviours occur at finite N: for example $g^{(2)}(0) = \frac{6}{5}(N-1)(N+3)/N(N+2)$ for finite N as $|g| \to \infty (|\Omega| \to \infty)$ and $g^{(2)}(0) \to 1.2$ as $N \to \infty$ in agreement with (6.21b). Notice too that for $N = 1$, $G^{(2)}(0) = 0$. This means the opposite of black-body coherence $g^{(2)}(0) = 2$. In the latter photons act as bosons and 'bunch'. Bunching leads to the observable intensity-intensity correlation $g^{(2)}(0) = 2$. Evidently $g^{(2)}(0) = 1$ means 'no bunching'. So $g^{(2)}(0) = 0$ means 'antibunching' — a feature associated with the fermion character of the electrons in atoms. Antibunching has been seen[43]. Some results for $G^{(1)}(0)$ calculated for different N from the present model are shown in Fig. 18(a). Results for $g^{(2)}(0)$ are shown in Fig. 18(b).

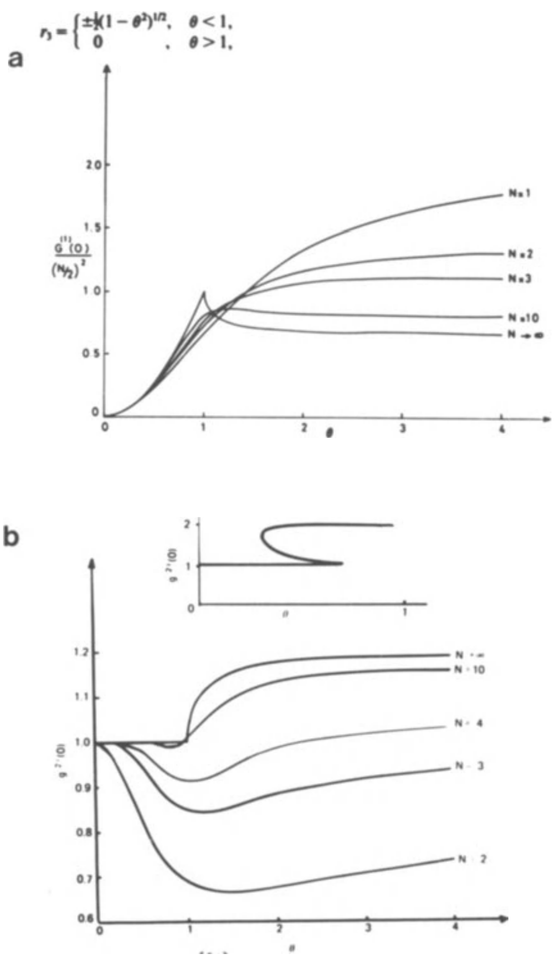


Fig. 18. (a) Results for $G^{(1)}(0) = \langle S_+ S_- \rangle$ calculated for different N; $g^{(2)}(0) \equiv \langle S_+^2 S_-^2 \rangle / |G_1(0)|^2$ is plotted in (b). Note that $g^{(2)}(0) \to 1.2$ (partial coherence) for $N \to \infty$ and $\Theta \to \infty$, and there is a sharp transition at $\Theta = 1$. For small finite N, $g^{(2)}(0) < 1$ (antibunching). The inset gives comparable results for the de-correlation (6.8).

These results of the exact quantum analysis tell us that a phase transition occurs in the coherently driven Dicke model which is not a phase transition of optical bistability type treated in §5. The de-correlation (6.8) leads to the optical bistability of (6.10) solely as a consequence of the approximation. There is a physical interpretation, because the radiation damping from simple atoms is isolated by the $-\frac{1}{2}\langle S_x \rangle$ term in (6.8) for example. The exact quantum theory contains this damping as part of the total collective action. In order to find optical bistability from the quantum model one must include (in some way) the effects of a cavity. We then find optical bistability from the coherently driven Dicke model providing the system is *not* decorrelated. Some results are shown in the Fig. 19. Our conclusion from this analysis is that both the quantum theory, and its fluctuations, *and* the cavity are necessary for the optical bistability phase transition. One or other by themselves is not sufficient. This conclusion seems to put in doubt all of the work of §§3-5 where neither a cavity or quantum features were taken into account! This point is yet to be understood. There seems to be a crucial difference between the small system (the point system of the Dicke model) and large systems. On the other hand the small sample Dicke model is NOT unphysical.

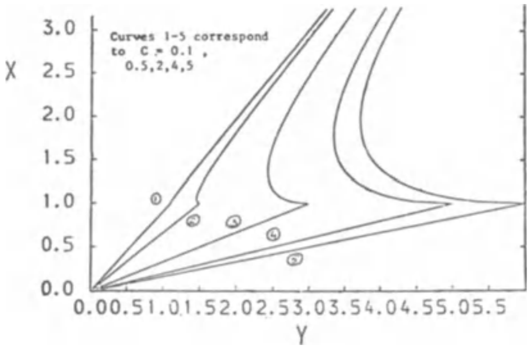


Fig. 19. Optical bistability for the quantum Dicke model coupled to a cavity. There is no optical bistability without the cavity or after the de-correlation approximation (that is without quantum fluctuations).

In this case we couple the model not to a single mode coherent driving field but to a broad band incoherent one. A typical and realisable example of a broad band chaotic field is a thermal field, that is black-body radiation. The equation one has to solve is the master equation[44,45].

$$
\begin{aligned}
\frac{d\rho}{dt} &= -i[(H_o+H_1+H_{af}),\rho] + \Lambda_a\rho + \Lambda_f\rho \\
H_o+H_1 &= \omega_s S_z + \omega_s a^\dagger a \quad ; \quad H_{af} = -ig(aS_+ - a^\dagger S_-) \\
\Lambda_a\rho &\equiv 2\gamma_o(1+\bar{n})[2S_-\rho S_+ - S_+S_-\rho - \rho S_+S_-] \\
&\quad + 2\gamma_o\bar{n}[2S_+\rho S_- - S_-S_+\rho - \rho S_-S_+] \\
\Lambda_f\rho &\equiv \kappa(\bar{n}+1)(2a\rho a^\dagger - a^\dagger a\rho - \rho a^\dagger a) + \kappa\bar{n}(2a^\dagger\rho a - aa^\dagger\rho - \rho aa^\dagger)
\end{aligned}
\tag{6.22}
$$

The operators a, $a^\dagger$ annihilate and create photons in a single cavity mode; $\Lambda_a\rho$ describes stimulated and spontaneous processes out of the upper atomic states (term in $(1+\bar{n})$) and stimulated processes into these (term in $\bar{n}$). The number $\bar{n}$ is the mean photon occupation number in the incoherent driving field of those modes resonant with the atoms at the frequency ω_s. For so called 'bad cavities' on which experiments have recently been done[37] it turns out that $\Lambda_a\rho$ can actually be dropped from (6.22).

If one uses $\langle S_z\rangle = \mathrm{Tr}(\rho S_z)$ and defines the atomic inversion $r_z \equiv N^{-1}\langle S_z\rangle$ much as above one finds (for very bad cavities $\kappa \to \infty$) that (6.22) reduces to the Einstein rate equation

$$\dot{r}_z = -\gamma_o(r_z+\tfrac{1}{2}) - 2\gamma_o\bar{n}\,r_z \tag{6.23}$$

if, but only if, all collective action between the atoms is ignored. This way one finds

$$r_z = -\tfrac{1}{2}/(1+2\bar{n}) \quad ; \quad \langle S_z\rangle = Nr_z = -\tfrac{1}{2}N/(1+2\bar{n})\,. \tag{6.24}$$

If N_+ atoms are in their upper states and N_- in their lower states evidently $N_+ + N_- = N$ and $N_+ - N_- = 2Nr_z$. Then $N_+ = \tfrac{1}{2}N + Nr_z = N\bar{n}/(1+2\bar{n})$. This result is essentially due to Einstein.

Actually the time constant in (6.23) is $\gamma_o + g^2\kappa^{-1}$, γ_o or $g^2\kappa^{-1}$ according as both of Λ_a and Λ_f or Λ_a or Λ_f alone are retained in (6.22). A very different result from (4.24) is obtained if all the collective effects are kept and one works with the full quantum theory (6.22) with or without Λ_a. One finds

$$N_+ = X[1 - (N+1)X^N + NX^{N+1}]/[1 - X - X^{N+1} + X^{N+2}] \tag{6.25}$$

where $X \equiv \bar{n}/(1+\bar{n})$. From (6.25) one finds

$$N_+ = \bar{n} / (1+2\bar{n}) \qquad N = 1$$
$$\rightarrow \bar{n} \text{ (simply) for } N \rightarrow \infty \quad . \qquad (6.26)$$

That this result is correct has just been shown by Haroche et al in the first of two experiments[37]. These scientists couple black-body radiation in a cavity to a beam of Na atoms which have been excited by a laser to the high Rydberg state $30S_{\frac{1}{2}}$. The cavity is tuned to $30S_{\frac{1}{2}} \rightarrow 30P_{\frac{1}{2}}$ transitions and the wave-length $\lambda \sim 2.25$ mm. This means there can be a large number of Na atoms (up to 30,000 in the experiments) within a volume $\sim \lambda^3$, so the Dicke model becomes applicable. In practice behaviours with asymptotes $2\bar{n}$ were seen (there are two possible polarisations). The Fig. 20 shows results for two different temperatures $T = 300^0$K ($\bar{n} = 47$) and $T = 900^0$K ($\bar{n} = 137$). The ordinate $\langle Ne(\Delta t)\rangle$ is what we have called N_+. It evolves to (6.25) and depends on time t .

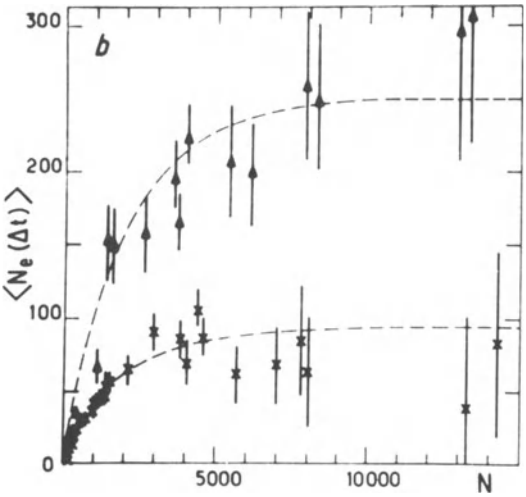


Fig. 20. Plots of the number of Na atoms $N_+ \equiv \langle Ne(\Delta t)\rangle$ in the Rydberg state $30P_{\frac{1}{2}}$ after starting in the $30S_{\frac{1}{2}}$ state. The atoms interact with a resonant mode of black body radiation at two different temperatures inside a tuned cavity. The asymptotes $2\bar{n} = 94$ ($T = 300^0$K) and $2\bar{n} = 274$ ($T = 900^0$K) of the solid curves form a good fit with the observed data (taken from S. Haroche in "Proc. Les Houches Summer School XXXVIII", R. Stora and G. Grynberg eds., North Holland, Amsterdam (to appear 1983)).

There is actually a very simple explanation of this behaviour for large enough N. The N 2-level atoms form a ladder of states separated by equal energies $\hbar\omega_s$. Together ther is the eigen-spectrum $n\hbar\omega_s$, $n = 1, 2, \cdots$ . . This is just the spectrum for what is a giant quantum oscillator. It will obey Bose-Einstein statistics, and the number of excited states is the number of excited atoms N_+ which is the number of photons excited at that frequency — namely $\bar{n}$. This is what is observed. Notice how the collective properties of the atoms have changed from those of a single atom!

In a second experiment[37] Haroche and colleagues observed the probability $P_m(t)$ of finding $m = N_-$ atoms in their ground states $28P_{\frac{1}{2}}$ at time t after starting at $t = 0$ in the excited states $29S_{\frac{1}{2}}$. It is known that collective action in this case leads to the emission of a short relatively intense pulse of light much more quickly than the spontaneous emission of radiation from the atoms acting independently. The phenomenon is called 'super-radiance' from the original work of Dicke who studied his model for initially inverted atoms without driving field[38]. It may also be called 'superfluorescence'[5]. A key feature is that the radiative time constant $\gamma_o = T_2^{-1}$ gains a collective factor $\sim N : \gamma_o \to N\gamma_o$.

Haroche's second experiment shows that $P_m(t)$ evolves through a Bose-Einstein type probability distribution, where it is monotonically decreasing with increasing m, through a bell-shaped region, and then finally to Bose-Einstein in equilibrium with the radiation. Some results are shown in Fig. 21: the solid curve is an approximate

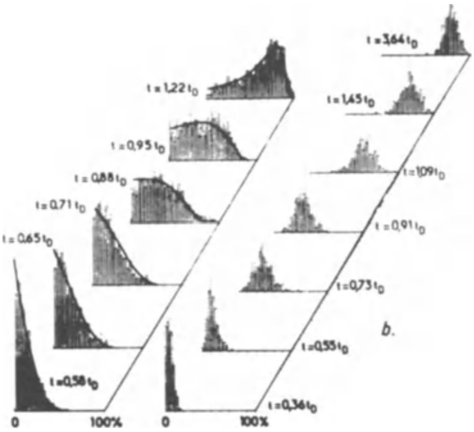


Fig. 21. Evolution of the probability $P_m(t)$ of finding $m = N_-$ atoms in their ground states ($28P_{\frac{1}{2}}$ states) after all starting from $29S_{\frac{1}{2}}$ at $t = 0$. The solid line is an approximate analytical solution of (6.22). The function $P_m(t)$ is of Bose-Einstein form for small and large t. But during super-radiant pulse emission $P_m(t)$ takes on the bell-shaped form typified by the laser in a coherent state.

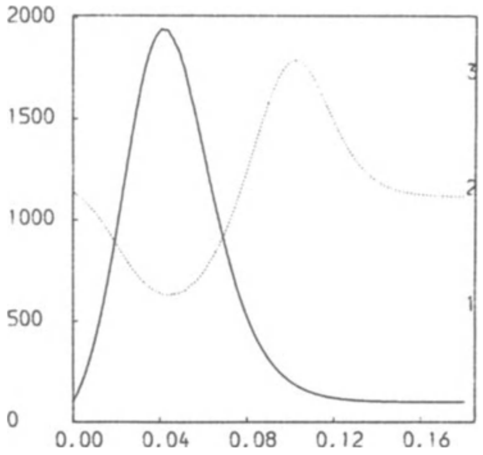


Fig. 22. Plot of $g^{(2)}(0)$ eq.(6.22)(dotted curve) at different times. Also shown, emitted super-radiant pulse (solid curve) $g^{(2)}(0) \approx 1$ (coherence) during the super-radiant emission, but settles on 2 after large swings): $g^{(2)}(0)=2$ is typical of incoherence and photon bunching.

analytical solution of (6.22). The bell-shaped region can be seen (but not the subsequent evolution to Bose-Einstein). This bell-shaped region has some aspects of the Poisson statistics associated with the coherent state[3] which is predicted for the single mode quantum laser well above threshold[3]. Thus it seems some aspects of coherence surround the dynamical process of super-radiant emission from initially inverted atoms. The final transition is to incoherence as with the state associated with the early Bose-Einstein type distribution. The final Fig. 22 shows a plot of $g^{(2)}(0)$ calculated from (6.22) at small times $t \ll \infty$: on the same Fig. 22 is imposed the profile of the super-radiant pulse intensity. Evidently coherence $g^{(2)}(0) \approx 1$ occurs during the maximum of the super-radiant emission and incoherence $g^{(2)} \approx 2$ is associated with the early and final states of the system.

7. CONCLUSION

It is possible to show from the accepted principles of quantum electrodynamics that a coupled system of 2-level atoms and quantised electromagnetic field is governed by the operator Bloch equation (2.3) defined at each point $\underset{\sim}{x}$ with $E(\underset{\sim}{x},t)$ now a quantised e.m. field operator together with the quantised form of the Maxwell equation (2.5). There are subtleties about the demonstration many of which are associated with radiative damping and radiative level shifts[46]. Neverthless one can derive damped *operator* Bloch-Maxwell equations this way with $T_2^{-1} \equiv \gamma_0$ $T_1^{-1} = 2\gamma_0$. Thus, of the quantities introduced in this paper only the cavity damping constant κ needs to be added phenomenologically. The work of §6 shows that the passage to the c-number Bloch-Maxwell equations (2.7), which are fundamental to this paper, will eliminate some specific quantum features. Nevertheless we have seen from the material presented in the paper that one set of equations, the Bloch-Maxwell equations, operator or c-number, describe a remarkably wide range of different observable optical phenomena. In particular we have seen that the Bloch-Maxwell equations in envelope form, which are a simple but, nevertheless intrinsically non-linear, system of equations, lead to quite different behaviours according as time scales, or space scales, or pumping parameters or coherence features are changed. For example in §3 evolution in space as well as in time is important; but the life times of the optical pulses involved means that both T_2 and T_1 can be dropped and κ becomes irrelevant. In §§4 and 5 spatial variation is factored out through cavity modes or spatial averages but T_2, T_1 and κ as well as pumping parameters become crucial to the behaviours. In §6, a quantum analysis, spatial variation is eliminated because the space scale is so small; but T_2 and T_1 play a role or not depending on interpretation and κ must be introduced for bad cavities.

This wide variety of behaviour becomes possible through the collective action of the atoms. Analogies of the subtle collective actions we have described may find an equally subtle role in biological systems perhaps not necessarily through the intermediary of the transverse part of the electromagnetic field handled in this paper. The remarkable dynamics displayed by the Lorenz equation forms (4.22) of the envelope Bloch-Maxwell equations emphasises the extraordinarily complicated oscillatory motions which can emerge from a relatively simple system of equations. They also illustrate how the same equations can emerge from two quite different physical contexts — laser physics on the one hand and convective hydrodynamical instability on the other. Complicated collective oscillatory motions must inevitably be a feature of biological systems and their existence will be important to their dynamical stability and flexibility of operation. We can only speculate whether the strange attractor plays any role in this dynamics.

In my view the ideas surrounding coherence and incoherence are likely to be just as important. It should be remarked here that, in the case of the coherently driven Dicke model of §6, the state with $<S_z> \neq 0$ for $0 < \Theta < 1$ is essentially an *atomic* coherent state. Such states are described[42] for all N although they become eigenstates of some operator for only $N \to \infty$. It seems more than possible that atomic coherent states such as these, as well as the many different dynamically controlled processes we have described in this paper, could all play their roles in biological systems.

The two most important points we have tried to stress in this paper are these: nonlinearities transform the actions of a system into a rich and various field of behaviours very different from those predicted by their approximated linearised forms. These behaviours are various and different even though the equations describing them remain basically the same. Key effects can be produced by the operation of very different time scales, or space scales, or both, or as bifurcations induced by what may be very small changes in control parameters. Collective action in which atoms co-operate to produce behaviours very different from those exhibited by single atoms is also a hall mark of all the processes we have described. Collective action, dynamically controlled, is surely the secret of all biological processes.

REFERENCES

1. W. P. Su, J. R. Schrieffer and A. J. Heeger, Phys. Rev. Lett. 42:1698 (1979); Phys. Rev. B22:2099 (1982). Also D. K. Campbell and A. R. Bishop, Nuclear Phys. B200 [FS4]:297 (1982).
2. W. P. Su, "Nonlinear Excitations in Polyacetylene", in: these Proceedings (1983); A. R. Bishop, "Solitons in Biological and Synthetic Polymers", in: these Proceedings (1983).
3. M. Sargent III, M. O. Scully and W. E. Lamb Jr., "Laser Physics", Addison Wesley Publ. Co., Reading, Mass. (1974).
4. R. K. Bullough, P. J. Caudrey and H. M. Gibbs,"The double sine-Gordon equations: a physically applicable system of equations", in: "Solitons" Topics in Current Physics 17, R. K. Bullough and P. J. Caudrey, eds., Springer-Verlag, Heidelberg (1980). Chap III.
5. R. K. Bullough, P. M. Jack, P. W. Kitchenside and R. Saunders, Physica Scripta 20:364 (1979).
6. R. K. Bullough, R. Saunders and C. Feuillade, "Theory of Far Infra Red Superfluorescence", in: "Coherence and Quantum Optics IV", L. Mandel and E. Wolf, eds., Plenum Press, New York (1978) p. 263.
7. R. K. Bullough and R. Saunders, "Theory of FIR Super-fluorescence", in: "Cooperative Effects in Matter and Radiation", C. M. Bowden, D. W. Howgate and H. R. Robl, eds., Plenum Press, New York (1977).
8. "Solitons" Topics in Current Physics 17, R. K. Bullough and P. J. Caudrey, eds., Springer-Verlag, Heidelberg (1980).
9. A. S. Davydov, Theor. and Math. Phys. 40:408 (1979); in: "Biology and Quantum Mechanics", Pergamon Press, Oxford (1982) p. 229; and "Energy Transfer Along Alpha-Helical Proteins" in: "New Horizons of Quantum Chemistry", D. Reidel Publ. Co., Dordrecht, Holland (1983) pp. 65-75.
10. A. C. Scott, "Solitons and Bioenergetics", in: these Proceedings (1983); Phys. Rev. A26:578 (1982).
11. R. K. Bullough, Phys. Bulletin Feb. (1978) pp. 78-82; "Solitons", in: "Interaction of Radiation with Condensed Matter, Vol. I", International Atomic Energy Agency, Vienna (1977) pp. 381-469.

12. W. R. Adey, "Nonlinear, Nonequilibrium Aspects of Electromagnetic Field Interactions at Cell Membranes" in: these Proceedings (1983).
13. H. Fröhlich, "General Theory of Coherent Excitations in Biological Systems: in: these Proceedings (1983).
14. J. Timonen and R. K. Bullough, Phys. Lett. A82A:183 (1981).
15. R. E. Shusher and H. M. Gibbs, Phys. Rev. A5:1634 (1972).
16. S. V. Manakov and I. R. Gabitov, Phys. Rev. Lett. 50:495 (1983).
17. S. V. Manakov, A. P. Fordy and R. K. Bullough, Phys. Letts. 91A:98 (1982).
18. R. Bonifacio and L. A. Lugiato, "Cooperative Effects in Optical Bistability and Resonance Flourescence", in: "Coherence and Quantum Optics IV", L. Mandel and E. Wolf, eds., Plenum Press, New York (1978) p. 249 and references.
19. L. D. Landau and E. M. Lifshiftz, "Statistical Physics", Pergamon Press, London (1959) p. 430.
20. P. Eugene Stanley, "Introduction to Phase Transitions and Critical Phenomena", Oxford Univ. Press, Oxford (1971) p. 14. See especially the Table 1.1 on p. 14.
21. G. S. Agarwal, Phys. News 13:1 (1982) Much of the discussion in §4 is based on this paper.
22. H. Haken, Phys. Lett. A53:77 (1975).
23. E. N. Lorenz, J. Atmos. Sci. 20:130 (1963).
24. H. L. Swinney and J. P. Gollub, "Hydrodynamic Instabilities and the Transition to Turbulence", Topics in Applied Physics 45, Swinney and Gollub (eds), Springer-Verlag, Heidelberg (1981).
25. H. Haken in his introduction to "Synergetics", H. Haken ed., Springer-Verlag, Heidelberg (1977), pp. 1-9.
26. P. Ruelle and F. Takens, Commun. Math. Phys. 20:167 (1971).
27. C. Sparrow, "The Lorenz equations, bifurcations and strange attractors", Springer Applied Mathematics Series 41, Springer-Verlag, Heidelberg (1982).
28. M. J. Feigenbaum, J. Stat. Phys. 19:25 (1978); Phys. Lett. A74: 375 (1979); Los Alamos Science 1:4 (1980).
29. L. N. Da Costa, E. Knobloch and N. O. Weiss, J. Fluid Mech. 109:25 (1981).
30. G. S. Agarwal, L. M. Narducci, D. H. Feng and R. Gilmore, "Dynamical Approach to Steady State and Fluctuations in Optically Bistable Systems", in: "Coherence and Quantum Optics IV", L. Mandel and E. Wolf, eds., Plenum, New York (1978) p. 281.
31. R. K. Bullough, S. S. Hassan and S. P. Tewari, "Refractive index theory of optical bistability", in: "Quantum Electronics and Electro-Optics", P. L. Knight, ed., John Wiley and Sons, Ltd., Chichester-New York (1983).
32. K. Ikeda, Opt. Comm. 30:257 (1979); K. Ikeda, H. Daido and O. Akimoto, Phys. Rev. Lett. 45:709 (1980).
33. W. J. Firth, Optics Comm. 39:343 (1981); E. Abraham, W. J. Firth and J. Carr, Phys. Letts 91A:47 (1982).

34. H. M. Gibbs, F. A. Hopf, D. L. Kaplan, and R. L. Shoemaker, Phys. Rev. Lett. 46:474 (1981); F. A. Hopf, D. L. Kaplan, H. M. Gibbs and R. L. Shoemaker, Phys. Rev. A25:2172 (1982).
35. D. W. McLaughlin, "Solitary Waves as Fixed Points of Infinite Dimensional Maps", in: these Proceedings (1983).
36. E. Del Giudice, "Order and Structures in Living Systems", in: these Proceedings (1983).
37. J. M. Raimond, P. Goy, M. Gross, C. Fabre and S. Haroche, Phys. Rev. Lett. 49:117 (1982); 49:1924 (1982).
38. R. H. Dicke, Phys. Rev. 93:99 (1954).
39. S. V. Lawande, R. R. Puri and S. S. Hassan, J. Phys B: Atomic and Mol. Phys. 14:4171 (1981); P. D. Drummond and S. S. Hassan, Phys. Rev. A22:622 (1980).
40. S. S. Hassan, R. K. Bullough, R. R. Puri and S. V. Lawande, Physica 103A:213 (1980).
41. R. R. Puri and S. V. Lawande, Phys. Lett. 72A:200 (1979); Physica 101A:599 (1980).
42. P. Carruthers and M. M. Nieto, Rev. Mod. Phys. 40:411 (1968); F. T. Arecchi, E. Courtens, R. Gilmore and H. Thomas, Phys. Rev. A6:2211 (1972).
43. H. J. Kimble, M. Dagenais and L. Mandel Phys. Rev. Lett. 39: 691 (1977); E. Jakeman, E. R. Pike, P. N. Pusey and J. M. Vaughan, J. Phys. A: Math. Gen. 10:L257 (1977); H. J. Carmichael, P. Drummond, P. Meystre and D. F. Walls J. Phys. A: Math. Gen. 11:L121 (1978); R. K. Bullough, in: the discussion to Phil. Trans. Roy. Soc. A 293:232-236 (1979).
44. S. S. Hassan, G. P. Hildred, R. R. Puri and R. K. Bullough, J. Phys. B: Atom. and Mol. Phys. 15:2635 (1982).
45. R. R. Puri, G. P. Hildred, S. S. Hassan and R. K. Bullough, "Exact Quantum Theory for N Rydberg Atoms in a Cavity", in: "Coherence and Quantum Optics 5", L. Mandel and E. Wolf, eds., Plenum Press, New York (1983).
46. F. Hynne and R. K. Bullough, "Quantum theory of multilevel many atom systems". To be published.

ENTRAINMENT-QUASIPERIODICITY-CHAOS-COLLAPSE: BIFURCATION ROUTES OF EXTERNALLY DRIVEN SELF-SUSTAINED OSCILLATING SYSTEMS

Friedemann Kaiser

Institute of Theoretical Physics, University of Stuttgart

7000 Stuttgart 80, FRG

ABSTRACT

Some theoretical models for nonlinearly oscillating phenomena in biological systems are discussed. We start with the necessary requirement for the creation of self-sustained oscillations (self-organization). Regular external perturbations interact with the internal oscillations, the system can be synchronized to the external drives (entrainment). A sharp frequency response results, exhibiting both, frequency and intensity windows and a rather irregular behaviour near the entrainment region.

A further increase of the external (static and periodic) fields can lead to sequences of period-doubling bifurcations, alternating with quasiperiodic and irregular regions (quasiperiodicity, chaos). As a consequence, a regularly driven self-oscillating system exhibits intrinsic chaotic behaviour though the underlying dynamic is strongly deterministic.

Finally, the energy input destabilizes the system (collapse), leading to the onset of propagating pulses (solitary waves), i.e. a nonlinear temporal structure gets replaced by a non-linear spatiotemporal structure.

The possible relevance for specific biological effects will be discussed, including threshold behaviour, time of irradiation, extreme sensitivity and its changes by external means.

INTRODUCTION

Fröhlich has promoted the ideas that coherent excitations can be established in metabolically active biological systems (1). He proposed three types of excitations within his general theoretical developments:

(1) high excitation of a single mode (Bose Condensation Model)
(2) excitation of highly polar metastable states (High Polarization Model)
(3) excitation of limit cycle oscillations (Coherent Oscillation Model).

The concept of coherent excitations has gained increasing support both by a series of theoretical investigations and by some extraordinary experimental results. A detailed discussion of Fröhlich's ideas as well as fascinating experimental results may be found in refs. (2, 3), where in addition much of the relevant literature can be found.

The different types of excitations should play an important role in both, the dynamic order and the specific function of biological systems. Consequently, one has to expect that excitations of the proposed types could possibly lead to changes in the behaviour and function of biosystems.

In the present work we restrict ourselves to the excitation of coherent oscillations of limit cycle type. Based on this concept some physical and mathematical models have been developed. In the "Brain Wave Model" (4) long range collective interactions within the Greater Membrane of the brain can establish slow chemical oscillations which are connected to low frequency electric oscillations by means of the large dipole moment of activated enzymes. ELF fields can interact with the internal limit cycle (coherent oscillation); its frequency and field dependent collapse is discussed. Thus an explanation of the extreme high sensitivity of certain biological systems to very weak electromagnetic signals is given: the external stimulus may only serve as a trigger to start an internal response signal.

The modelling of coherent oscillations by limit cycles (self-sustained oscillations) is intriguing. Besides hysteresis behaviour, i.e. bistability, this type of dissipative structure in systems far from thermodynamic equilibrium seems to play a dominant role in chemical, biochemical and biological systems.

An extension of a simple limit cycle system which has to include at least two stable limit cycles has been presented to model specific brain experiments (5). The external perturbation of limit cycle type of oscillations is the main topic of our presentation, the emphasis of which is the different bifurcation routes from the onset of the oscillation to its collapse.

We do not intend to discuss the detailed absorption mechanism of electromagnetic energy by biological materials. Some comments on the related questions can be found in ref. (5), where both, the underlying dielectric properties and the consequences for extremely weak field interactions are considered.

Furthermore, a number of contributions within this volume yield more information on both, the theoretical background and the experimental findings (vid. the papers of Fröhlich, Schwan, Illinger and Keilmann, Webb, Adey, to mention a few).

I PHASE A: ONSET OF A LIMIT CYCLE

A limit cycle is a self-sustained oscillation: at the expense of a non-periodic internal source of energy a periodic process is created and stabilized. It should be emphasized that necessary requirements for a limit cycle to be built up are*

- nonlinear interactions and processes
- dissipative processes

Without any external drive, a nonlinear oscillation can result from self-organizing processes, i.e. an intrinsic cooperative behaviour of the system dominates, no externally forced order is required. Both, the amplitude X_{LC} and the frequency ω_{LC} of the limit cycles are completely determined by internal parameters.

---

* Limit cycles seem to be an adequate description of periodic processes and they are of fundamental importance in most modelling approaches. Solitons on the other hand are of increasing importance in the description of propagation processes. However, they describe processes in conservative systems, where limit cycles are not possible. Since real systems are dissipative, loss processes should be included. These would lead to a soliton-like behaviour, i.e. solitary wave propagation and other travelling waves. Though of fundamental different origin, solitons (dispersive, nonlinear) and solitary waves (dissipative, nonlinear) are connected when only the resulting processes are regarded: signals are propagated losses are of secondary importance within the time of propagation. From this point of view, solitons seem to be able to play a dominant role for specific transport mechanisms similar to limit cycles for periodic processes.

Concept

Periodic internal processes are modelled by limit cycle oscillations. The underlying mathematical description is given by nonlinear differential equations, the solutions of which are stable periodic attractors (limit cycles): LC (X_{LC}, ω_{LC}).

Models

M I: generalized Van der Pol oscillator

$$\boxed{\begin{aligned} \dot{x} &= y \\ \dot{y} &= \mu\,(1 - x^2 + \alpha\, x^4 - \beta\, x^6)\, y - x \end{aligned}} \tag{1}$$

This set of equations can exhibit 2 stable limit cycles with an unstable one in between. x is the amplitude ofoscillations, e.g. a measure for the polarization; α, β, $\mu > 0$ are parameters. The essential terms are the nonlinear and linear damping terms (dissipation), the competition of which is responsible for the onset of limit cycles (5, 6).

M II: coherent oscillation model (Brain wave model)

$$\boxed{\begin{aligned} \dot{\nu} &= \gamma\,\sigma + \alpha\, A\, \sigma\, \nu + (c^2\, e^{-\Gamma^2 \nu^2} - d^2)\, \nu \\ \dot{\sigma} &= -\beta\nu - \alpha\, A\, \sigma\, \nu \end{aligned}} \tag{2}$$

This model is a combination of both, a nonlinear chemical reaction and specific dielectric processes (ferroelectric tendency). It combines slow chemical and electrical processes. Whereas the chemical terms ($\gamma\sigma$, $\beta\nu$, $\alpha\, A\, \sigma\, \nu$) describe a slow Lotka-Volterra dynamics with an underlying very fast process (hidden in the parameters α, β, γ), the ferroelectric term exhibits a sort of Van der Pol process: competition of nonlinear gain and linear loss processes.

The model has been proposed by Fröhlich (4) on the basis of his general concept in order to describe slow oscillations in the greater membrane of brain. $N = \nu + \gamma/\alpha A$ and $S = \sigma + \beta/\alpha A$ are the concentrations of active enzyme and of substrate molecules, respectively. All parameters are positive. For $c^2 > d^2$ a stable limit cycle exists; furthermore, an unstable steady state solution (fixpoint) or two unstable fixpoints and a stable one exist. The limit cycle exhibits the slow chemical and electrical oscillations of the model, both of which are coupled by the highly polar state of the enzyme. The resulting oscillation and phase plane diagrams are shown in Fig. 1.

Model II has been investigated in detail for a time-independent external perturbation (constant electric fields). The possible results are given in refs. (7, 8). For our purposes it suffices to mention a specific result: if the external field strength F_l goes beyond a critical threshold value $F_{l,c}$, the oscillating state bifurcates into a stable, non oscillating one (the internal oscillation is frozen in).

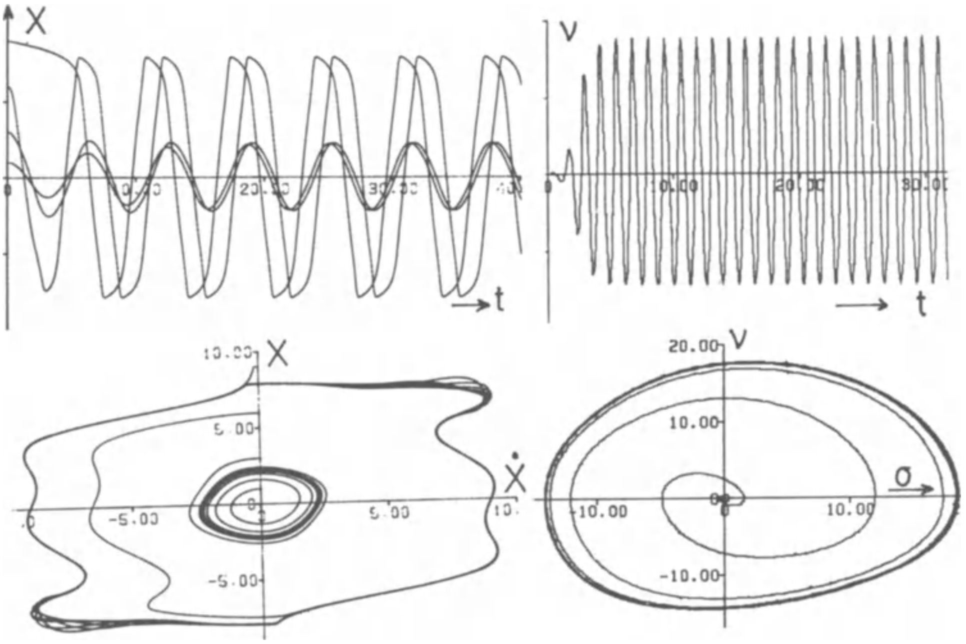


Figure 1 : Oscillation ($x - t$; $V - t$) and phase plane ($x \dot{-} x$; $V - \sigma$) diagrams of MI (l.h.s.) and MII (r.h.s.) ,respectively. For MI four different initial conditions are taken.

II PHASE B: PERTURBATION OF A LIMIT CYCLE

The central point of the coherent excitation concept is the perturbation of the internal oscillation(type 3 of excitations):

$$L\ C\ (X_{LC},\ \omega_{LC}) + F(t) \longrightarrow\ ?$$

We restrict ourselves to a systematic external influence; neither, internal nor stochastic perturbations are discussed. Quite generally, we deal with systems of the following type:

$$\begin{aligned} \dot{x} &= P\,(x,\, y) + F_o \cos \lambda\, t \\ \dot{y} &= Q\,(x,\, y) \end{aligned} \qquad (3)$$

or equivalently

$$\begin{aligned} \dot{x} &= P\,(x,\, y) + F_o \cos \phi \\ \dot{y} &= Q\,(x,\, y) \\ \dot{\phi} &= \lambda \end{aligned} \qquad (4)$$

System (3) is a nonautonomous 2-d system, system (4) the equivalent 3-d autonomous system. Eqs. (1), (2) are supplemented by a forcing term (vid. eqs. 3) to account for an external drive. A competing process results between the free internal and the periodic external oscillation.

II.1. Phase B, Step 1: Entrainment

An increasing external field strength F_o drives the free oscillation into a synchronized or entrained state, if $F_o > F_{o,E}$:

$$L\,C\,(X_{LC},\, \omega_{LC}) \xrightarrow[F_o \cos \lambda t]{} LC\,(X_\lambda,\, \lambda)$$

The system obeys the external drive, it oscillates with the external frequency and with an amplitude depending on F_o. In addition, the oscillating system is driven through a series of quasi-periodic and nearly nonperiodic intermediate states. On the route from the free to the forced oscillation,one finds near the region of entrainment a nonlinear superposition of free and forced oscillations:

free internal oscillation $\longrightarrow$ "mixture" of free and forced oscil. $\xrightarrow[F_{o,E}]{}$ forced oscillation

The mathematical results and related specific effects are shown in the following scheme:

frequency response curve: sharp resonances complete entrainment, $F_o > F_{o,E}$: threshold long transient periods: time of irradiation effects		MI, MII
transitions $LC_1 \leftrightarrow LC_2 \leftrightarrow$ $\rightarrow$ nonoscillating states	frequency windows intensity windows	MI

These results are the typical ones for externally driven self-sustained oscillating systems. A well-known example is the Van der Pol oscillator (9). For MI and MII, details of the transitions to entrainment are to be found in refs. (3,5,6). Furthermore, some other specific types of behaviour have been investigated, the most relevant of which are briefly discussed:

1. for $\lambda \ll \omega_{LC}$ on finds a partial occurrence of the fast internal oscillation during one period of external drive (i.e. $2\pi/\lambda$) in the case of MI (fig. 3),

2. for MII the fast internal oscillation (frequency conditions as in case 1)is dominating, the slow external drive leads to an increasing modulation of the amplitude with the external frequency (beat phenomenon) (fig.2),

3. MI exhibits,in addition, for certain values of F_o and λ transitions from LC_1 to LC_2 and vice versa. Details depend on the initial conditions. For both transitions,the small limit cycle LC_1 is in its entrained state, whereas the large amplitude cycle LC_2 is in its free internal state. In addition, large transients do exist. For a rather restricted frequency region,only LC_1 exists for F_o near $F_{o,c}$ (specific intensity window).

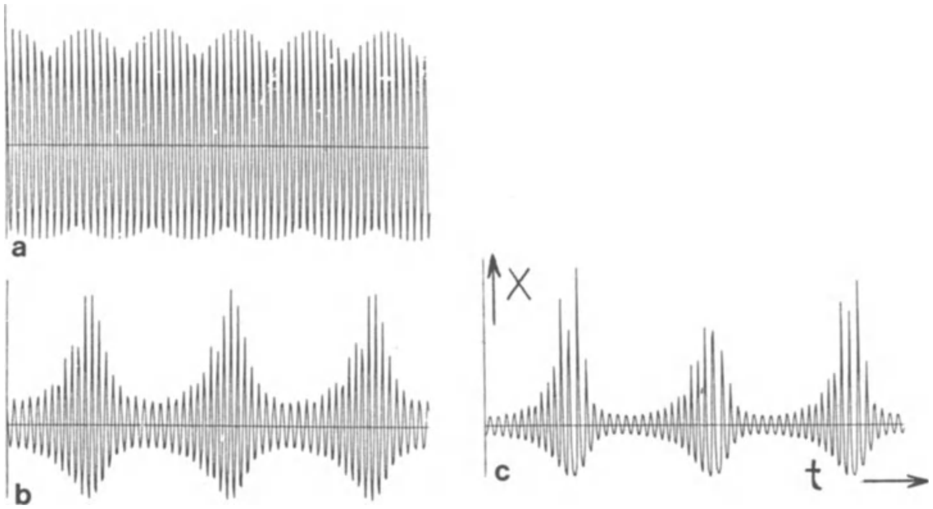


Figure 2 : Oscillation diagrams for MII with $\lambda \ll \omega$. Compared to MI (vid.fig. 3), the behaviour is completely different. No entrainment can be reached with increasing F_o, the beat oscillations collapse before the system is entrained.

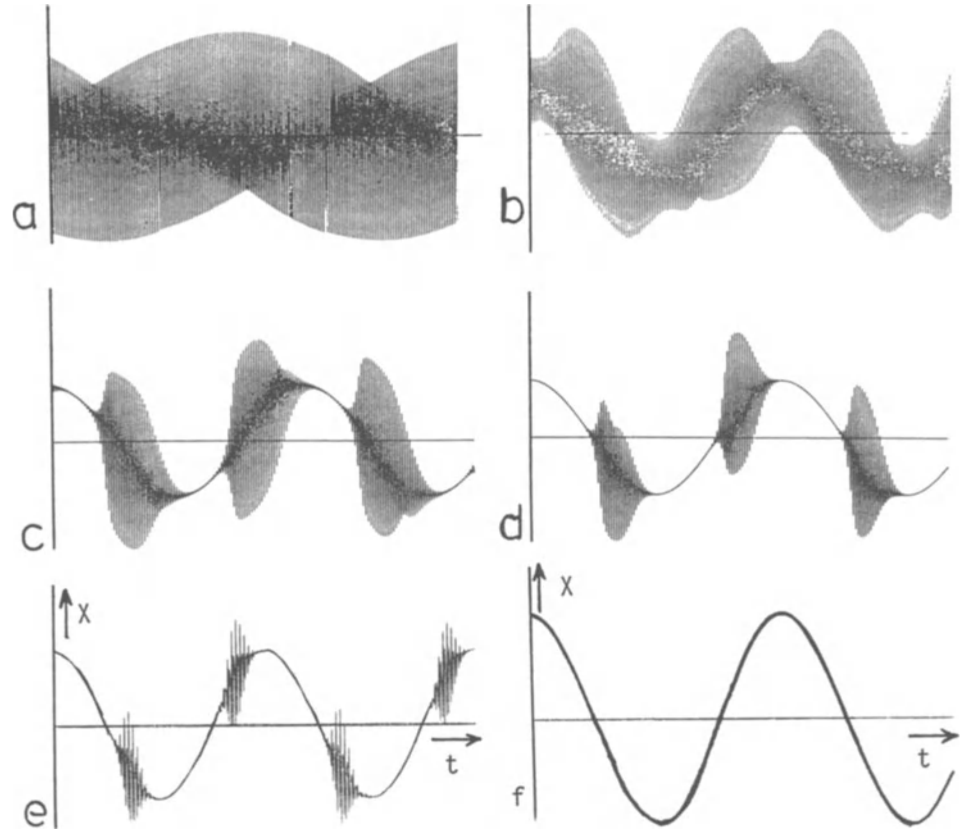


Figure 3 : Oscillation diagrams for MI with $\lambda \ll \omega$. F_o increases from a) to f), the transition from the fast free to the slow entrained oscillation is clearly exhibited.

II.2. Phase B, Step 2: Collapse

An essential part of our modelling is aimed at an investigation of the excitation of the entrained internal oscillation and its subsequent collapse leading to the onset of a response signal:

$$LC\ (X_{LC},\omega_{LC}) \longrightarrow LC\ (\lambda) \longrightarrow \text{Collapse}$$

MI allows for these steps only for the small amplitude oscillation LC_1, where in most situations one finds:

$$LC_1\ (\lambda) \longrightarrow LC_2\ (\tilde{\omega}, \tilde{\hat{x}})$$

$$LC_2\ (\tilde{\omega}, \tilde{\hat{x}}) \longrightarrow LC_2\ (\lambda)$$

The large amplitude oscillation remains stable. MII, on the contrary, always performs a transition to a collapsing state. The critical field strength $F_{o,c}$ is strongly dependent on the external frequency (fig. 4).

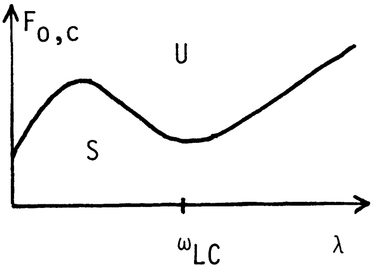


Figure 4 : Critical field strength $F_{o,c}$ as a function of the external frequency λ (MII). Above the line $F_{o,c}(\lambda)$ the system is unstable. U : unstable, S : stable region.

II.3. Phase B, Step 3: Onset of Pulse Propagation

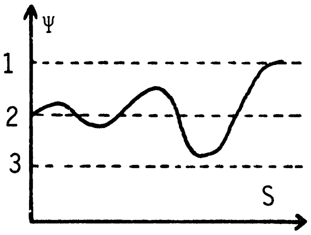
The collapse of the oscillating system (MII) leads to the onset of specific types of travelling waves and propagating signals. A detailed investigation of the completely time-dependent problem requires rather complicated manipulations, which are still in progress.

A study of the problem, where only a constant external field and diffusive processes as the only space-dependent terms are applied, has led to some interesting results (10):

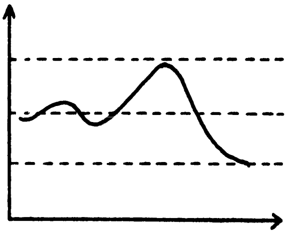
1. the limit cycle remains stable with respect to spatial perturbations for $F_1 < F_{1,c}$;
2. for $F_1 = F_{1,c}$ (i.e.bifurcation of the spatially homogeneous problem to a stable nonperiodic attractor) a spatial perturbation destabilizes all spatially homogeneous solutions;
3. for $F_1 > F_{1,c}$ no temporally or spatially stable solutions exist.

A detailed stability and phase plane analysis shows, that three types of travelling waves do exist. The stability of these pulses depends on the internal parameters and on the resulting velocity of propagation.

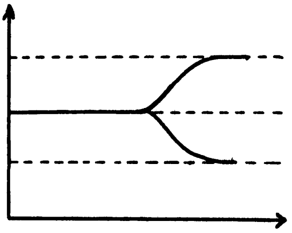
Schematically, the different types of pulse propagation exhibit the following transitions:



Solitary wave
(transition like)



Solitary wave
(transition like)



Transition wave
(up+down transition)

III DIFFERENT ROUTES TO COLLAPSE

The transition from the entrained to the collapsing states of M II yields bifurcation routes to excitation (11). The combination of a time-dependent and a constant field (i.e. $F(t) = F_o \cos\lambda t + F_1$) reveals further bifurcations.

Separately, two processes govern different behaviour:

$F(t) = F_1$: periodic attractor $\xrightarrow[F_{1,c}]{}$ static attractor

$F(t) = F_o \cos\lambda t$: periodic attractor $\xrightarrow[F_{o,c}]{}$ collapse

A superposition of both, the static and the periodic field is more complicated:

$F(t) = F_o \cos\lambda t + F_1$: competition of both processes.

The results are shown in fig.5 .

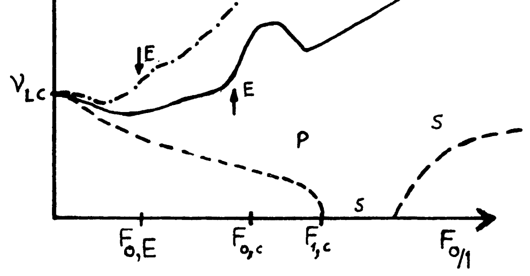


Figure 5 **:Steady state amplitude ν_S of MII as a function of the external field strength.**
----- $F(t) = F_1$
-·-·- $F(t) = F_o \cos\lambda\ t$
—— $F(t) = F_o \cos\lambda\ t + F_1$
(E= entrainment, P= periodic state, S= static state, * = collapse)

It should be emphasized that the application of a constant field leads to a drastic enlargement of the stability region, i.e. the limit cycle collapse is shifted to much larger field strengths F_o (see fig. 6).

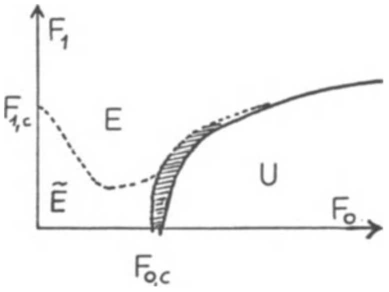
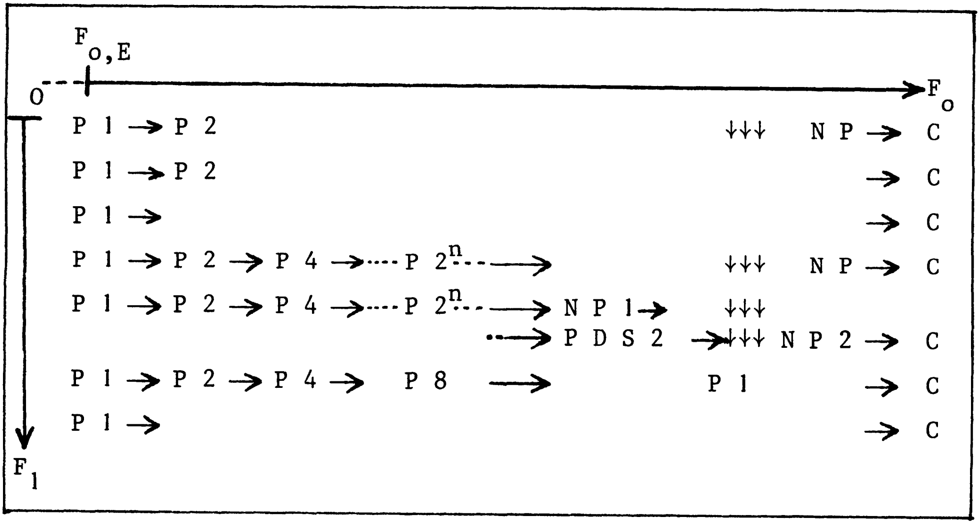


Figure 6 : Stability region, given by the external fields F_o and F_1 . Frequency λ is fixed, MII. E: complete entrainment, $\tilde{E}$: free + forced oscillations, U : unstable, ▨:bifurcations.

The shift of the collapse, caused by F_o and F_1, respectively, is accompanied by a series of bifurcations within this region of regained stability. These series are shown in the following scheme as a function of F_o and F_1 (λ is fixed):



P 1 → P 2 ↓↓↓ N P → C
P 1 → P 2 → C
P 1 → → C
P 1 → P 2 → P 4 →····P 2^n···→ ↓↓↓ N P → C
P 1 → P 2 → P 4 →····P 2^n···→ N P 1 → ↓↓↓
··→ P D S 2 → ↓↓↓ N P 2 → C
P 1 → P 2 → P 4 → P 8 → P 1 → C
P 1 → → C

P 1 = L C (λ) ; P 2 = L C ($\lambda/2$) ; P 2^n = L C (λ/n) ;
N P 1,N P 2 = nonperiodic, irregular, chaotic motion;
C = collapse ; ↓↓↓ onset of noise ;
P D S 1 = period doubling sequence (P1,P2,P4,P8,...P2^n,...);
P D S 2 = period doubling sequence (P6,P12,P24,P48,..P$3\cdot2^n$,...).

The above given bifurcation scheme is only a small part of the whole route of an internal oscillation under the influence of an external drive. The complete route can briefly be described by the sequence:

$$LC(\omega_o) \longrightarrow P/NP/QP \longrightarrow LC(\lambda\ ) \longrightarrow \{scheme\} \longrightarrow C$$

free oscil.	intermed. states	entrained oscil.	bifurc. routes	collapse

For an increasing periodic field strength F_o,at least six types of bifurcations occur from entrainment (E) to collapse (C):

1. E → no bifurc./no noise → C
2. E → 1 PD - bifurc./no noise → C
3. E → 1 PD - bifurc./ noise → chaos → C
4. E → 4 PD - bifurc./no noise → P 1 → C
5. E → ∞ PD - bifurc./ noise → chaos → C
6. E → ∞ PD - bifurc./ noise → chaos →
→ PD - bifurc./ noise/periodic windows → C

(PD = period doubling bifurcation)

An interesting feature is the occurrence of strongly periodic windows within the chaotic regions of type six of bifurcations.Furthermore, one finds transitions with and without period doubling bifurcations. In addition, some of the processes are dominated by an increase of the noise level and an irregular motion, whereas some others (F_o small and F_o large, respectively)run into their collapse without any noise or chaotic motion (type 1,2,4).

In figure 7 we have drawn the maximum amplitude of oscillations the limit cycle can have in its entrained state (i.e. P1),where F_o is that critical field strength above which a bifurcation to P2 or to some other states occurs. Two regions with respect to F_o exist, where only P1 - states (oscillations with the external frequency) are present.

Three different processes compete (or cooperate in some sense) before the collapse occurs:

| period doubling
| amplitude increase/decrease
| noise increase

One of these processes domitates; in addition,the whole transitions can be suppressed by a large increase of beat modes (fig. 2).

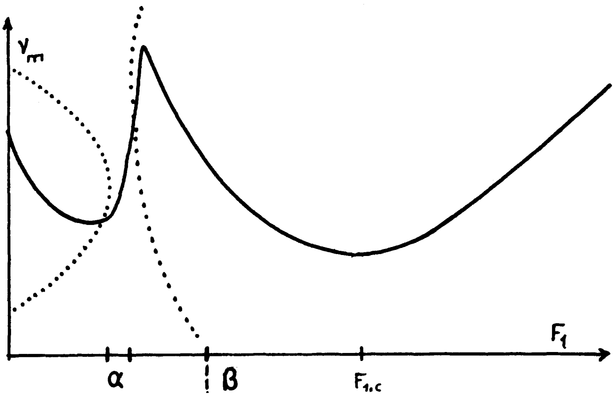


Figure 7 : Maximum amplitude V_m as a function of the static field F_1. For F_o, the maximum field strength has been taken that is possible to keep the P1 - state stable. The dotted line shows the same situation for the state P2 (no P2 - and higher states exist in the regions α and β).

A more detailed view on the bifurcation route is given in a series of pictures for a constant static field F_1 and for a certain external frequency. The following four pictures belong to the type 6 of the bifurcation routes.

Fig. 8 shows the power spectra (Fourier transformed of the amplitude of oscillations) for several periodic and nonperiodic states of oscillations,

Fig. 9 is a representation of some of the corresponding stropo-scopic phase portraits,

Fig. 10 is a drawing of the complete bifurcation diagram. Both, the strongly periodic and the irregular regions as a function of the external periodic field are exhibited.

Finally, in fig. 11 a logarithmic plot of the external field F_o demonstrates the width of the periodic and nonperiodic regions and windows.

A discussion of these results will be given in the next two chapters. Special emphasis is led onto both, a mathematical relation and a biological interpretation of the different types of motions.

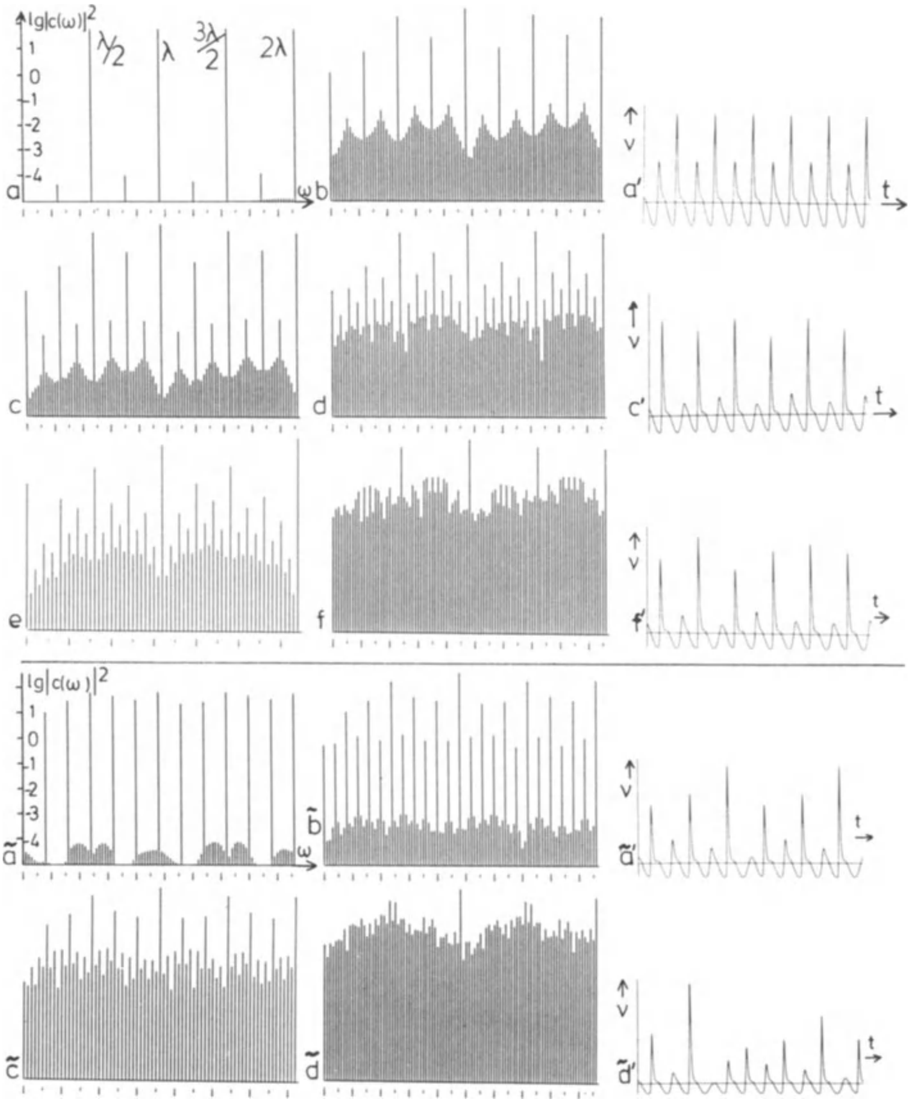


Figure 8 : Power-spectra ($\lg c|(\omega)|^2$ -ω -diagrams) for λ and F_1 fixed (MII). F_0 increases from a) to f) and then from ã) to d̃). The whole series shows the type 6 bifurcation route from entrainment (a) to collapse (d̃). The figs.a',c',f',ã', and d̃' are the corresponding oscillation diagrams.
The whole series is :
P2,P4,P8,P16,P32,NP1,P6,P12,P18,NP2.

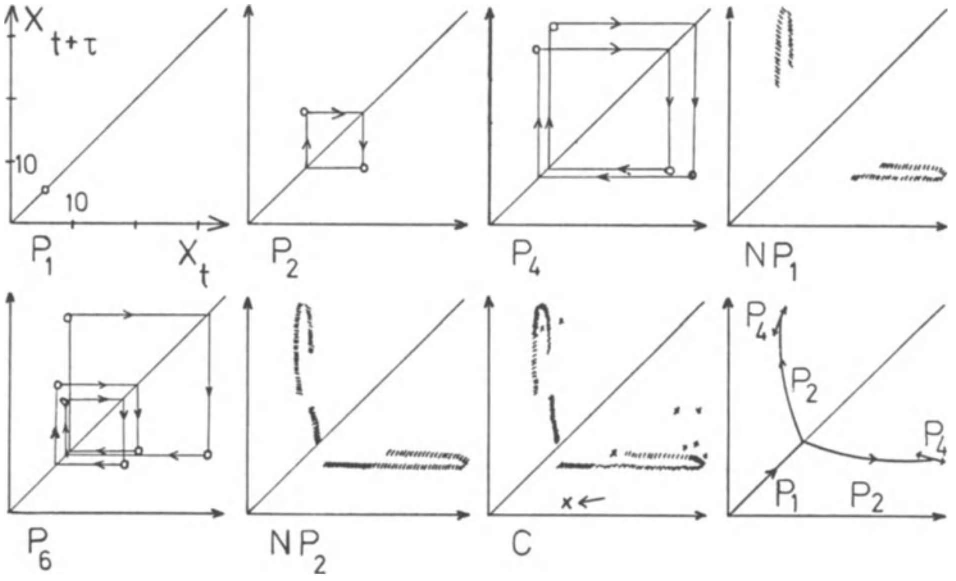


Figure 9 :Stroposcopic phase portraits ($x_{t+\tau}$ - x_t- diagrams, $\tau = 2\pi/\lambda$)for some of the states of fig. 8. The last diagram of the series shows the development of the amplitudes for the P1 - P2 - P4 bifurcation route.

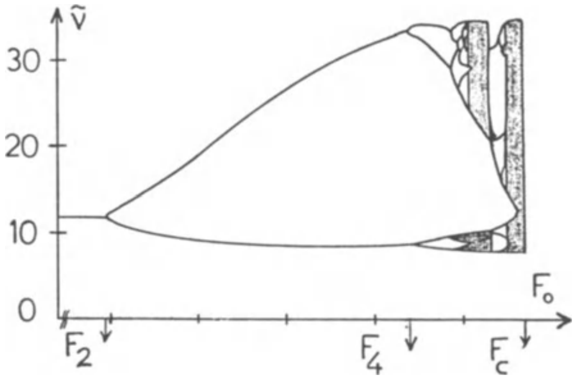


Figure 10:Bifurcation diagram (stroposcopic amplitude $\tilde{V} \equiv x_t$ as a function of F_0).F_1 and λ are fixed to the same values as in fig.8.
F_2 : bifurcation to P2 state,
F_4 : bifurcation to P_4 state.

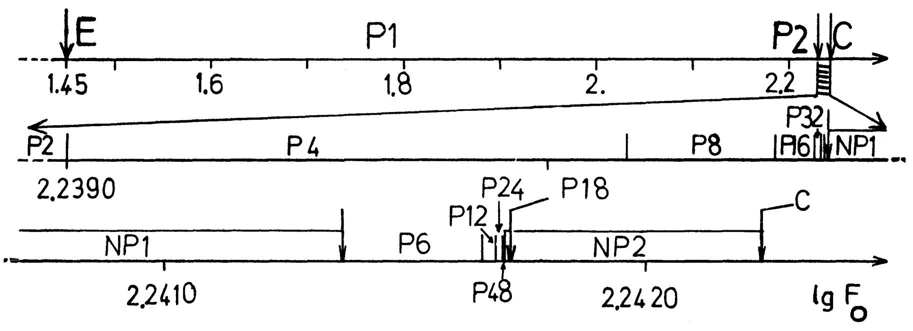


Figure 11 : Sequences of period doubling bifurcations, chaotic states and periodic windows for $F_{o,E} \leq F_o \leq F_{o,c}$**. The second and the third line are extensions of the shaded region shown as a small part of the top line. The external parameters** λ **and** F_1 **are fixed.**

IV SOME MATHEMATICAL ASPECTS

Model MII exhibits a bifurcation behaviour that, at least in its essential points, is wellknown from a number of dissipative dynamical systems (nonlinear differential equations and nonlinear maps). Examples cover nearly all domains of physics and related fields.Nonlinear optics, hydrodynamics, chemical kinetics and electrical circuits are some important areas of both, theoretical investigations and experimental studies. Either, they include the variation of an internal parameter (system parameter) or periodic and static external drives (see 12 - 14 for further references).

The essential points of model MII are as follows:

1. the existence of accumulation points and the subsequent onset of chaotic regions,
2. the occurrence of sequences of period doubling bifurcations,
3. the nonexistence of aperiodic states and of noise within one period-doubling sequence,
4. extremely sharp peaks,
5. the existence of periodic windows within the chaotic regions,
6. the onset of reverse bifurcations above and close to the points of accumulation (i.e. within irregular regions the lowest frequencies disappear subsequently).

All these features are characteristic ones for a number of nonlinear maps. Furthermore, MII seems to fullfill the scaling properties which have been found by Feigenbaum for a whole family of nonlinear maps.Finally, it should be emphasized that a chaotic behaviour is exhibited in completely deterministic systems (chaotic is taken as a synonym for irregular or nonperiodic or weakly turbulent motion).

The question "why do numerous nonlinear systems run through such a complicated series of bifurcations"? requires an amount of further investigations. We have started work on the different ways a driven system can dissipate the energy it receives by an external drive.Furthermore, a detailed discussion of the dynamics of bifurcations from a physical point of view is necessary. The results are in a very preliminary state and are omitted.

V CONSEQUENCES , BIOLOGICAL RELEVANCE?

We want to draw our attention to the consequences of our results. In addition,a possible relevance of the different steps and phases on the way from the free internal oscillation to the collapse will be discussed. We start with a further interpretation of the bifurcation routes:

Limit cycle : The system is in a self-organized state,i.e.a periodic pattern is developed and maintained by nonlinear processes and the influx and efflux of energy. A typical example are coherent excitations and oscillations and the related energy storage mechanism.

⇓ Application of fields/irradiated systems

Entrainment : The system is synchronized to the external drive. A number of specific effects occur, they offer a very first explanation of some specific effects found in biological irradiation experiments. We only mention the occurrence of sharp resonances, windows, time of irradiation and threshold effects, partial sensitivity losses and sensitivity increases. All these effects are weak intensity effects (nonthermal effects);they are far from being explained.

↓

Bifurcations: The functional order of the model systems is reorganized.The occurrence of strongly periodic states with submultiples of the external frequency enables the system to couple to other frequencies or to decouple partially from the driving field.

↓

Collapse : The onset of travelling waves causes the system to respond to extremely weak external signals,since most of the energy stems from the internal oscillation.A transmission of a variety of signals (messages) is encoded in the waveforms.

From our model calculations and some other work (e.g.15) we are able to derive some rules:

- both, periodic and stochastic perturbations can drive a limit cycle oscillation into a chaotic state (if certain requirements with respect to the field strength and frequencies are fullfilled)

- chaotic states are very sensitive to periodic perturbations, a periodic oscillation is regained

- chaotic states remain chaotic under stochastic influences (fluctuations, noise....).

The completely different behaviour of periodic and chaotic states under periodic drives offers some additional possibilities a biological system can respond to radiation. From this point of view, both, regular and irregular motion of brain waves should be of a certain importance for the function of the brain and for its response to external fields.However, there is still a very large way to be gone from simple models to the complicated biological systems. One possible route has been given by Fröhlich.It looks promising. Others might also be of importance.

REFERENCES

1. H.Fröhlich, "General theory of coherent excitations on biological systems" (this volume)
2. H.Fröhlich, The biolocical effects of MW and related questions", Advances in Electronics and Electron Physics, 53: 85-152 (1980)
3. F.Kaiser, "Nonlinear oscillations (limit cycles) in physical and biolocical systems", in "Nonlinear electromagnetics", P.L.E. Uslenghi ed., Academic Press, New York (1980)
4. H.Fröhlich, in "Brain interactions with weak electric and magnetic fields",S.M.Bawin and W.R.Adey eds., Neurosci.Res.Progr.Bull. 15: 67-72 (1977)
5. F.Kaiser, "Theory of resonant effects of rf and mw energy" in : "Biological effects and dosimetry of nonionizing radiation", S.Michelson ed., Plenum Press, New York ,251-282 (1982)
6. F.Kaiser, "Coherent modes in biolocical systems,perturbation by external fields", in : "biological effects of nonionizing radiation", K.Illinger ed.,ACS Symp.Ser. 157: 219-241 (1981)
7. F.Kaiser, "Limit cycle model for brain waves",Biol.Cybernet.27: 155 (1977)
8. F.Kaiser, "Coherent Oscillations in biological systems I ", Z.Naturforsch. 33a: 294 (1978)
9. B.Van der Pol, "Forced oscillations in a circuit with nonlinear resistance", Phil. Mag. Ser. 7, 13: 65 (1922)

10. F.Kaiser, "Coherent oscillations in biological systems II", Z.Naturforsch. 33a: 418 (1978)
11. F.Kaiser, "Specific effects in externally driven self-sustained oscillating biophysical model systems", in "Coherent excitations in biological systems", H.Fröhlich and F.Kremer eds.,Springer Verlag, Berlin (1983)
12. J.-P.Eckmann, "Roads to turbulence in dissipative dynamical systems", Rev.Mod.Phys. 53: 643 (1981)
13. E.Ott, "Strange attractors and chaotic motions of dynamical systems", Rev.Mod.Phys. 53: 655 (1981)
14. P.Collet,J.-P.Eckmann, "Iterated maps on the intervall as dynamical systems", Birkhäuser, Boston (1980)
15. G.Mayer-Kress,H.Haken, "Influence of noise on the logistic model", Journ.Stat.Phys. 26: 149 (1981)

FIELD INDUCED RESONANCE AND CHAOS IN NONLINEAR CONSERVATIVE SYSTEMS

L.E. Reichl

Center for Statistical Mechanics
University of Texas at Austin
Austin, TX 78712

ABSTRACT

Two models are discussed which show how dynamic external fields can induce internal chaos and enhance dissociation in nonlinear conservative systems.

INTRODUCTION

Until now there has been very little theoretical work which can clarify the process by which an external electromagnetic field interacts with the conservative nonlinear internal dynamics of molecules. The work that has been done can be used to describe this interaction only in the continuum limit of the molecule and does not shed much light on the mechanisms involved.(1) There has been a great deal of work on the external field--molecule interaction in which the internal dynamics of the molecule is approximated by a linear model. Linear models serve to locate some of the natural frequencies of the molecule, but they give a completely incorrect description of the internal dynamics of the molecule when a field is present.

There is another closely related problem of importance in biology and that has to do with the effect of dynamic external fields on the diffusion of particles across potential barriers (pore penetration). This problem has been studied extensively, when no external field is present, using stochastic methods.(2) But little work has been done when dynamic external fields are present.

One of the main difficulties in modeling the effects of dynamic external fields on nonlinear systems is to find nonlinear systems which are completely solvable when the field is absent. Solutions to the unperturbed problem can then be used to study the perturbed problem. In this paper we will describe the results of work we have done in determining the effect of a dynamic, single frequency, external field on two simple conservative nonlinear systems--the two particle Toda lattice and a particle trapped in a double well (quartic) potential. We will find that any intuition we had developed based on linear model is completely wrong. As we shall show, a dynamic field profoundly distorts and disrupts the phase space of a nonlinear conservative lattice or barrier system and can induce large energy fluctuations and chaos in regions where we do not expect it.

In Section II, we describe work done on the perturbed two particle Toda lattice by L.E. Reichl, R. de Fainchtein, T. Petrosky and W.M. Zheng(3,4). In Section III, we describe work done on the perturbed double well system by L. E. Reichl and W. M. Zheng.(5) In Section IV, we give some concluding remarks

II. TODA LATTICE

The Toda lattice consists of a chain of masses coupled by forces which vary exponentially as the spacing between masses is changed. The Toda potential energy may be written

$$V(q) = \frac{A}{B} e^{-Bq} + Aq, \quad AB>0, \tag{II.1}$$

where A and B are constants and q is the relative spacing of two masses. The Toda potential is compared to the harmonic potential and the Morse potential in Figure 1. It approximates fairly well the behavior of real molecules when their energy is not too high. The Toda lattice is the simplest of a class of non-linear lattices which are known to be integrable and which sustain soliton propagation. For the case of the Toda lattice, the solitons consist of localized regions of compression on the lattice (localized concentrations of energy) which can travel long distances without being dispersed. This type of soliton has been observed in computer studies of long one-dimensional molecules.(6) The continuum limit of the Toda lattice is described by the Korteweg-de Vries equation--the equation most commonly used to model soliton propagation in fluid systems.

We are interested in the detailed mechanism by which a dynamical external field might disrupt this system. Perturbation methods do exist in the literature for studying the effects of external fields in the continuum limit using inverse scattering theory. (1,7) We are currently involved in such studies. However, this approach obscures the detailed mechanisms involved in the field-system interaction. Therefore, we decided to study the effect of a dynamic external field on the discrete Toda lattice.

The simplest case we can consider is the two-particle Toda lattice in the presence of a dynamic external field.(3,4) We can apply an external field in such a way that the center of mass is not displaced. Then the Hamiltonian for the perturbed two-particle lattice takes the form

$$H = \frac{P^2}{m} + \frac{A}{B}\left(e^{-Bq} + e^{Bq} - 2\right) + \varepsilon q \cos(\Omega_t) \qquad \text{(II.2)}$$

where m is the mass of each particle on the lattice, p and q are the relative momentum and position, respectively, of the masses, ε is the amplitude of the external field, and $\Omega_t = \omega_0 t + \delta$, where ω_0 is the radial frequency of the external field and δ is its initial phase. We may perform a canonical transformation to new variables (J,θ) defined by

$$p = \sqrt{\frac{mA}{B}}\,\frac{2k}{\sqrt{1-k^2}}\,\mathrm{sn}\,(\theta,\ k) \qquad \text{(II.3)}$$

and

$$q = \frac{2}{B}\,\ell n\left[\frac{\mathrm{dn}(\theta,k) - k\,\mathrm{cn}\,(\theta,k)}{\sqrt{1-k^2}}\right] \qquad \text{(II.4)}$$

where sn, dn, and cn are Jacobi elliptic functions, k is the modulus,

$$k^2 = \frac{J^2 - 4m_0\,A/B}{J^2}, \qquad \text{(II.5)}$$

and $m_0 = 4m/B^2$. Note that physical values of J must satisfy the condition $J^2 \geqslant 4m_0\,A/B$. The canonical transformation (II.3) and (II.4) is independent of ε and holds even when $\varepsilon = 0$.

We may write the Hamiltonian in terms of these new canonical variables

$$H = \frac{J^2}{m_0} - \frac{4A}{B} + \frac{2\varepsilon}{B}\cos(\Omega_t)\,\ell n\left[\frac{\mathrm{dn}(\theta,k) - k\mathrm{cn}(\theta,k)}{\sqrt{1-k^2}}\right] \qquad \text{(II.6)}$$

or if we Fourier expand we obtain

$$H = \frac{J^2}{m_0} - \frac{4A}{B} + \varepsilon \sum_{\substack{n=-\infty \\ (\text{odd})}}^{\infty} g_n(J) \cos\left(\frac{n\pi\theta}{2K} - \Omega_t\right) \qquad (\text{II.7})$$

where

$$g_n(J) = -\left[\frac{nB}{2} \sin h \left(\frac{n\pi K'}{K} \right)\right]^{-1} \qquad (\text{II.8})$$

and $K=K(k^2)$ and $K' = K(1-k^2)$ are complete elliptic integrals of the first kind.

The Fourier expansion of q indicates that this system for $\varepsilon = 0$ has natural frequencies

$$\omega_n = \frac{nJ\pi}{m_0 K} \quad (n \text{ odd}). \qquad (\text{II.9})$$

In the limit of very low energy the frequency ω_1 approaches the harmonic limit

$$\omega_1 = \frac{J\pi}{m_0 K} \xrightarrow[K \to 0]{} \left[\frac{AB}{m}\right]^{\frac{1}{2}} \qquad (\text{II.10})$$

In Figure (2a), we give a plot of the (p,q) phase space trajectory of the unperturbed lattice ($\varepsilon = 0$) for J = 491 and constants A=0.15, B=3.1, and m= 11.68×10^5 in atomic units (we use these values of A, B, and m throughout this paper). This is a low energy, and the system is near the harmonic limit. In Figure (2b) we show the spectrum of this system. The spectrum clearly shows that the system is near the linear limit. The frequency n=1 contains the most energy. Frequency n=3 is the only other one with significant energy. Because the system is nonlinear, the natural frequencies vary as the energy of the lattice varies. For this reason, unlimited energy absorption of the type seen in (unphysical) linear systems cannot occur. The system can resonate with an external field at all its natural frequencies, but once energy is absorbed the natural frequencies will change and the system will be driven off-resonance.

In Eq. (II.4) we see evidence of these nonlinear resonances. Let us focus on the resonance with frequency ω_n and approximate the Hamiltonian by

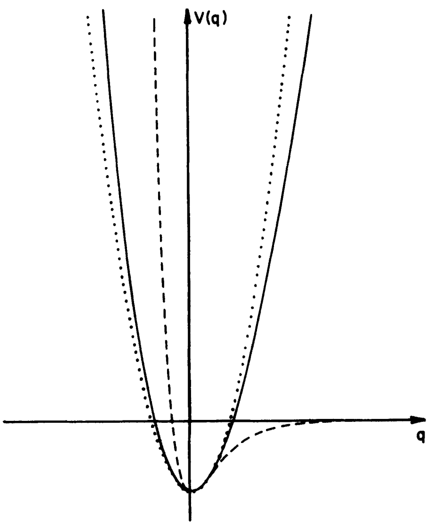


Figure 1: Comparison of the Toda potential (solid line), the harmonic potential (dotted line) and Morse potential (dashed line).

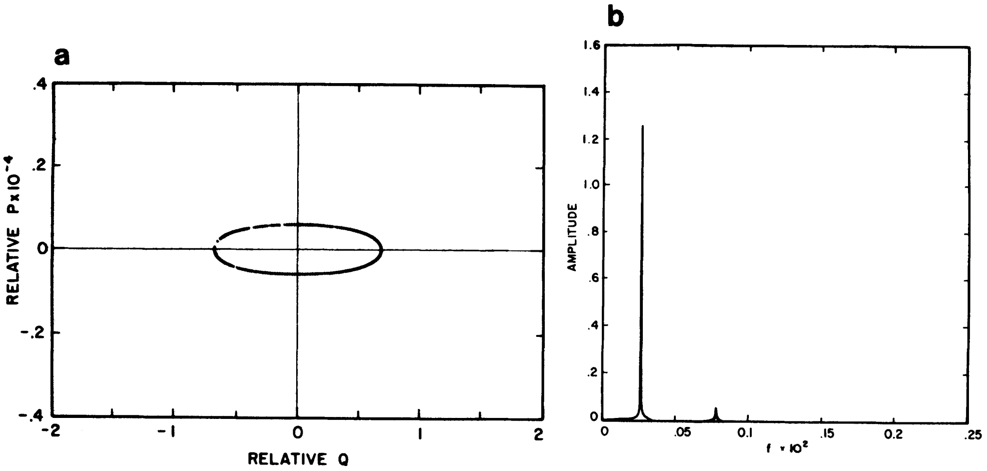


Figure 2: (a) The (p,q) phase space trajectory and (b) the spectrum of the unperturbed Toda lattice for $J = 491$, $A = 0.15$, $B = 3.1$ and $m = 11.68 \times 10^5$.

$$H \approx \frac{J^2}{m_0} - \frac{4A}{B} - \varepsilon |g_n| \cos\left(\frac{n\pi\theta}{2K} - \Omega_t\right) + \ldots \qquad \text{(II.11)}$$

This is basically the Hamiltonian of a pendulum and describes a localized disturbance in phase space which looks like the pendulum phase space. We can make a canonical transformation to a new set of coordinates (I,ψ) whose origin is located at $J = J^c$ (the center of the resonance zone) and which lies in the rest frame of the resonance zone. The generating function for this transformation can be written

$$F = -(\psi + \omega_0 t) \int_{J_c}^{J} \frac{2K(J')dJ'}{n\pi} \qquad \text{(II.12)}$$

The new coordinates are defined through the equations

$$I = -\frac{\partial F}{\partial \psi} = \int_{J_c}^{J} \frac{2K(J')dJ'}{n\pi} \approx \frac{2K(J_c)}{n\pi}(J - J_c) + \ldots \qquad \text{(II.13)}$$

and

$$\Theta = -\frac{\partial F}{\partial J} = (\psi + \omega_o t)\frac{2K(J)}{n\pi}, \qquad \text{(II.14)}$$

where we have expanded (II.13) in a Taylor series about J_c. The value of J_c is determined from the condition

$$\dot{\Phi} = \frac{2K(J_c)\,\omega_o}{n\pi} = \frac{2J_c}{m_o}. \qquad \text{(II.15)}$$

If k^2 is not close to one and if we stay in the disturbed region, we can terminate this Taylor series after the first term. In terms of this new coordinate system, the Hamiltonian takes the form

$$H' = H + \frac{\partial F}{\partial t} \approx \frac{n^2\pi^2 I^2}{4m_o K^2(J_c)} - \varepsilon |g_n| \cos\psi \qquad \text{(II.16)}$$

A phase space plot of this system is given in Figure (3). This is the basic shape of a resonance in a nonlinear system. The system resonates, changes energy, is driven off resonance, returns to its original state and begins the process again. Eq. (II.15) locates the positions of all the primary resonances.

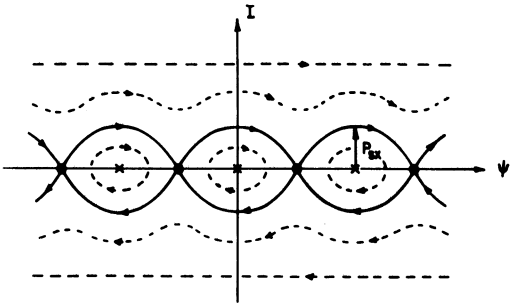


Figure 3: Sketch of the pendulum phase space. The solid line is the separatrix. Outside the separatrix we have rotation and inside oscillation. Dark circles are unstable (hyperbolic) fixed points. Crosses are stable fixed points. Dotted and solid lines correspond to possible trajectories of the pendulum

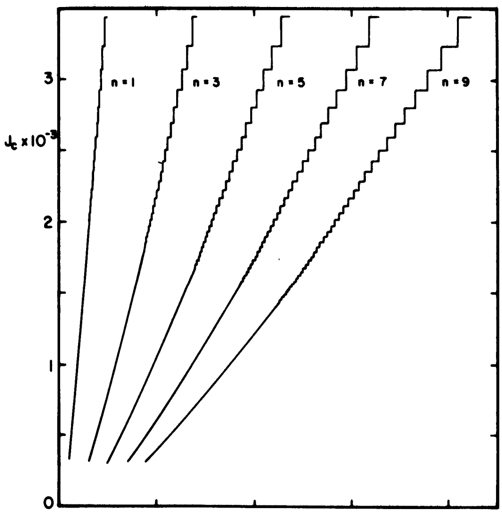


Figure 4: Plot of J_n^c vs f_0 for n = 1, 3, 5, 7 and 9 for the two particle Toda lattice with A= 0.15, B = 3.1, and m = 11.68 x 10^5.

It is useful to note that the system described by the Hamiltonian (II.7) is equivalent to a conservative two degree of freedom system with Hamiltonian

$$H = \frac{J_1^2}{m_o} - \frac{4A}{B} + J_2 + \varepsilon \sum_{\substack{n=-\infty \\ (\text{odd})}}^{\infty} g_n(J_1) \cos\left[\frac{n\pi\theta_1}{2K(J_1)} - \omega_o\theta_2 - \delta\right] \qquad \text{(II.17)}$$

Thus this new Hamiltonian, H, is a constant of the motion and is sometimes called the quasi-energy. The phase space of the perturbed system is four dimensional.

Let us now return to Hamiltonian (II.7). We see that primary resonance zones will be located in the phase space at values of J^c given by Eq. (II.15). A plot of the positions of the primary resonance zones as a function of frequency is given in Figure (4). No primary zones are expected below $f_0 = 2 \times 10^{-4}$. The primary zones emerge at $J^c=306$ and move to higher energies as the frequencies of the applied field are increased. We can probe the flow of points in a given region of phase space. We will start at $p = 600$ and $q = 0$, or equivalently $J = 491$, $\theta=K$. When $\varepsilon= 0$ the flow is given in Figure (2a). When $\varepsilon \neq 0$ the system has a four-dimensional phase space but we can obtain strobe plots of p and q. That is, we plot points p and q at time intervals $t = 1/f_0$ where $f_0 = \omega_0/2\pi$. As the frequency of the applied field increases, more and more resonance zones emerge in the phase space at $J = 306$ and then move to larger values of J as frequency increases. Thus, if we focus on point $J=491$, we will see resonance zones moving past from lower to higher values of J as frequency increases. In Figure (5), we show the approach and passage of the primary $n = 3$ resonance zone when $\varepsilon = 1.3$. This zone approaches from lower energy; we enter it at one of its hyperbolic fixed points. In Figure (5a), we are outside the resonance zone and the phase plot is distorted but qualitatively similar to the unperturbed case, although there is oscillation in energy. We enter the separatrix in Figure (5c), and by Figure (5e) it has passed.

In Figure (6) we show the approach and passage of the primary zone $n = 5$ when $\varepsilon= 1.3$. This sequence differs from that above in that we enter at the widest part. In Figure (5b) we are on the separatrix. In Figures (5c) and (5d) we are on inner islands, and in Figure (5f) we have passed out of the resonance zone. A theory which predicts the width of both of these zones is given in Reference 3.

If we examine resonance zone $n = 1$, it is not so simply described in terms of a simple primary resonance. Indeed, a whole sequence of resonances occur as we try to enter zone $n=1$. The sequence is shown in Figure (7) where $\varepsilon = 1.3$. The separatrix of

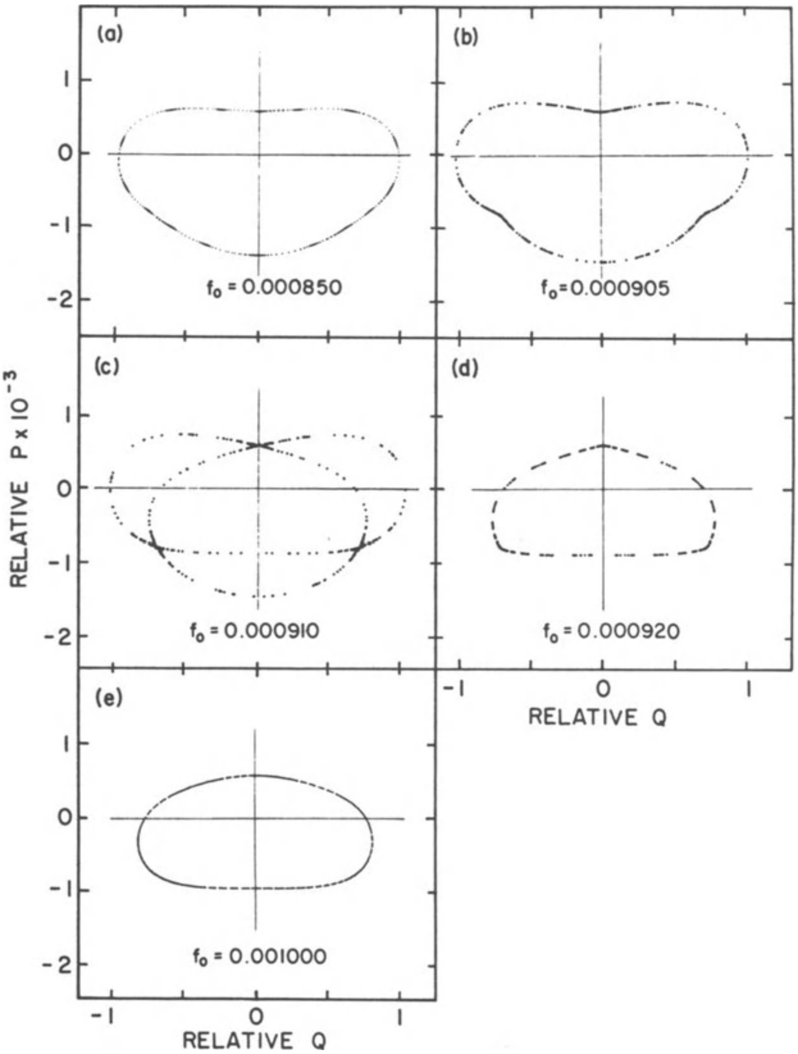


Figure 5: The primary n = 3 resonance zone for $\varepsilon = 1.3$.

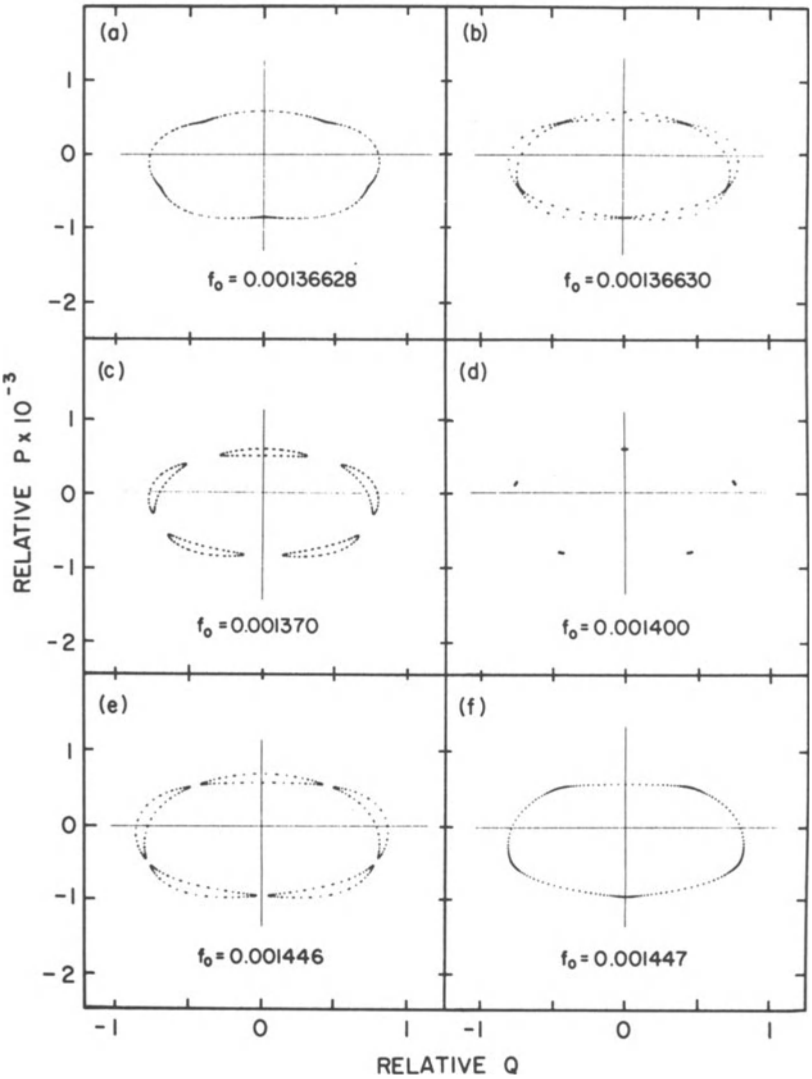


Figure 6: The primary n = 5 resonance zone for $\varepsilon = 1.3$.

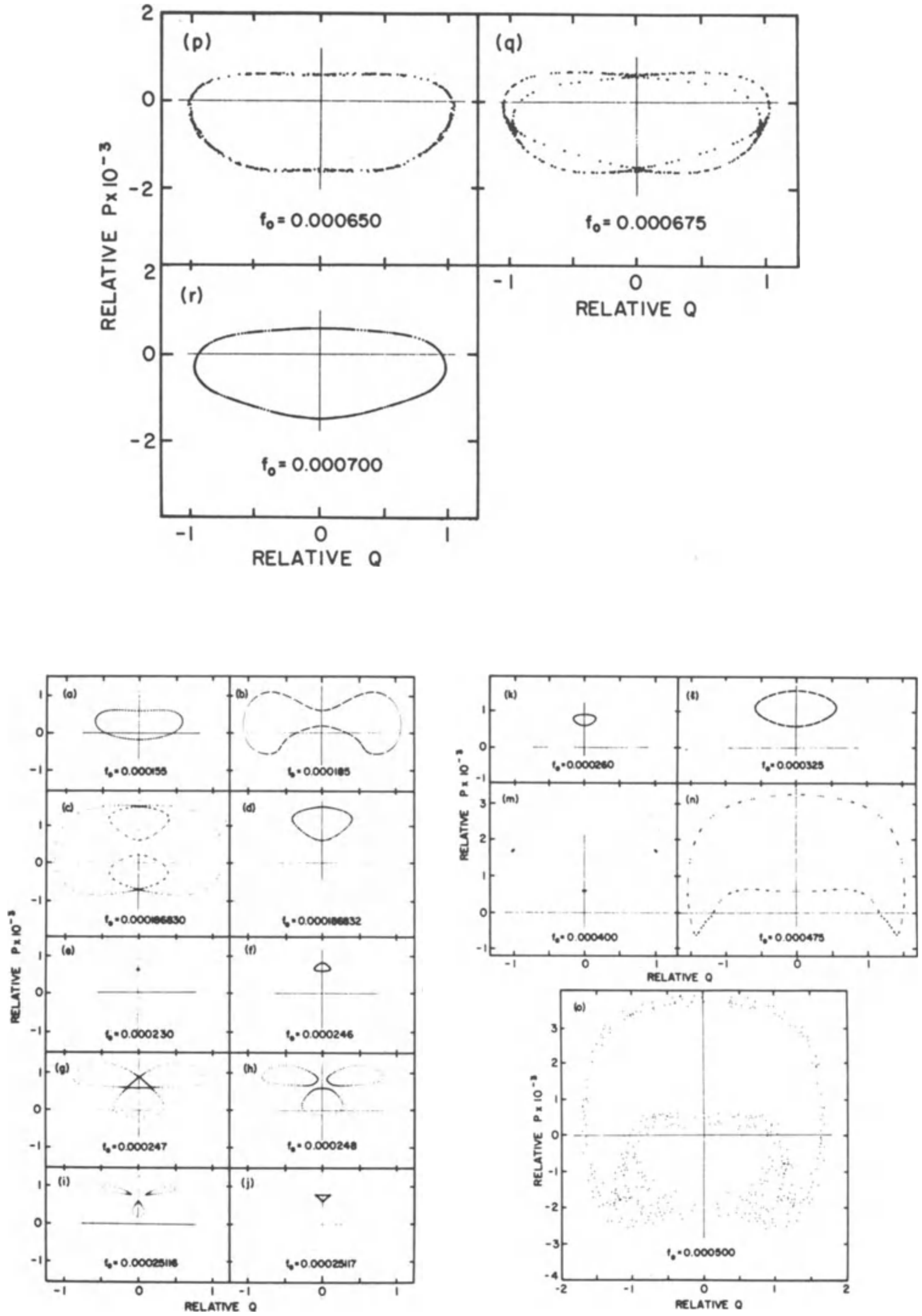


Figure 7: Resonance sequence leading into and out of the primary n=1 resonance zone for $\varepsilon = 1.3$.

the n=1 resonance curve appears in Figure (7.0) and is higly chaotic This chaos persists at lower amplitudes as well. Figure (8a) shows the separatrix of the n = 1 primary zone at $\varepsilon=0.8$, and in Figure (8b) we show its spectrum. We see that it too contains chaos and its spectrum has broadened.

What we are seeing in Figures (7) and (8) is the effect of so-called secondary resonance zones. The resonance zones induced by the external field actually form a dense set in the phase space. The primary zones (those seen directly from Eq. II.7) are the dominant resonance zones and serve to delineate the basic structure of the perturbed phase space, but they also comprise only a small fraction of the resonance zones that are actually induced by the external field. Between any two primary zones we can find an infinite family of secondary zones. Between any two secondary zones of a given family we can find still another infinite sub-family of secondary zones, etc. The stochastic region surrounding the n = 1 primary resonances is evidence that the secondary zones in the neighborhood of the n=1 primary resonance are overlapping and allow the system trajectory to wander somewhat chaotically through the phase space in that region.

One way to see that a dense set of resonance zones does indeed exist in the phase space (see Reference 3 for another view) is to focus on the region between two primary zones, say n_0 and n_0+1 in Hamiltonian (II.7). We then approximate the Hamiltonian by

$$H = \frac{J^2}{m_o} - \frac{4A}{B} + \varepsilon g_{n_o}(J)\cos\left(\frac{n_o\pi\theta}{2K(J)} - \Omega_t\right) + \varepsilon g_{n_o+1}(J)\cos\left(\frac{(n_o+1)\pi\theta - \Omega_t}{2K(J)}\right) \tag{II.18}$$

If we introduce the canonical transformation to new coordinates (I,ψ) whose origin is located at $J^c_{n_0}$ and lies in the rest frame of the $n_0{}^{th}$ resonance zone, then the Hamiltonian takes the form

$$H \approx \frac{n_o^2\pi^2 I^2}{4m_o K^2(J_c)} - \varepsilon|g_{n_o}|\cos\psi - \varepsilon|g_{n_o+1}|\cos\left(\kappa(\psi - \Delta vt)\right) \tag{II.19}$$

where $\Delta v = -\omega_0/n_0+1$ is the speed of zone (n_0+1) in (I,ψ) space and $\kappa = n_0+1/n_0$. By still another canonical transformation, namely,

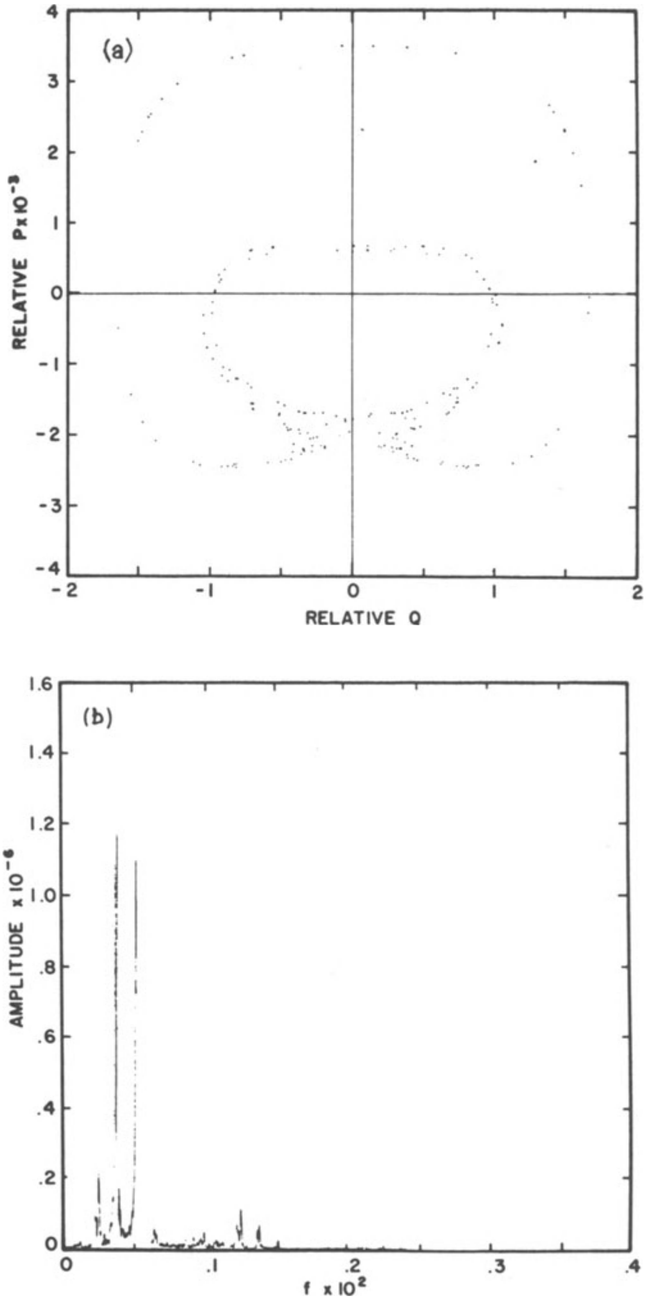


Figure 8: (a) Primary n = 1 resonance zone for $\varepsilon = 0.8$. (b) Spectrum of primary n = 1 resonance zone for $\varepsilon = 0.8$.

$$I = I_1 \,\mathrm{dn}\,(\frac{\psi_1}{2}) \quad \text{and} \quad \psi = 2\,\mathrm{am}\,(\frac{\psi_1}{2}) \qquad \text{(II.20)}$$

(am is the Jacobi amplitude function), we may write the Hamiltonian (II.19) in the form (5)

$$H = \frac{I_1^2}{2m_1} + \sum_{\ell=-\infty}^{\infty} V_{\ell_o}(I_1) \cos\left[(\kappa+\ell_o)\frac{\pi\psi_1}{2K(I_1)} - \kappa\Delta vt\right] \qquad \text{(II.21)}$$

with

$$V_{\ell_o}(I_1) = \frac{1}{4K} \int_{-2K}^{2K} d\psi_1 \cos\left[\kappa\,\psi(I,\psi_1) - (\kappa+\ell_o)\frac{\pi\psi_1}{2K}\right] \qquad \text{(II.22)}$$

Thus for the region between primary resonance zones n_0 and $n_0 + 1$ we again have an infinite number of resonance zones.

The effect of the various resonance zones on the physical behavior of the system is largely determined by the weight $g_{n_0}(J)$ associated with the primary resonance zones in addition to the shape of the spectrum of the unperturbed system. In the two-particle Toda system, the dominant disturbances are due to the n = 1 primary resonance zone and the secondary zones in its neighborhood. This is clearly seen in Figure (9) where we give a phase diagram of observed regions of chaos as a function of amplitude and frequency of the applied field. We see that for the two-particle Toda lattice the greatest disturbance occurs when the external field has a frequency slightly lower than or equal to the lowest natural frequency of the system. The induced chaos at frequencies below the lowest natural frequency of the system is unexpected.

The unperturbed two-particle Toda system is completely stable. If we wish to model the effect of external fields on pore penetration (diffusion over barriers) or dissociation, the simplest system to look at is that of a particle in a quartic potential.

III. THE QUARTIC POTENTIAL(5)

Let us now consider the motion of a particle in a quartic potential and in the presence of sinusoidally varying external field. The Hamiltonian for this system can be written.

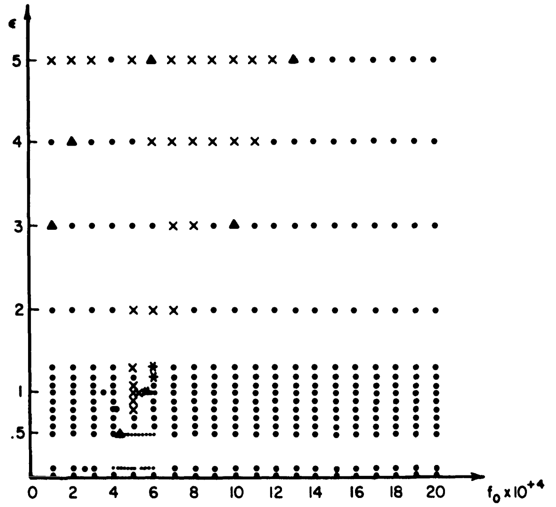


Figure 9: Phase diagram of regions of field induced chaos (as a function of amplitude ε and frequency f_0 of the applied field). Crosses show obvious chaos. Triangles show regions of marginal chaos. Stars show regions for which chaos lies between marginal and obvious. Dots show regions with no observable chaos. No strobe plots were obtained for unmarked regions.

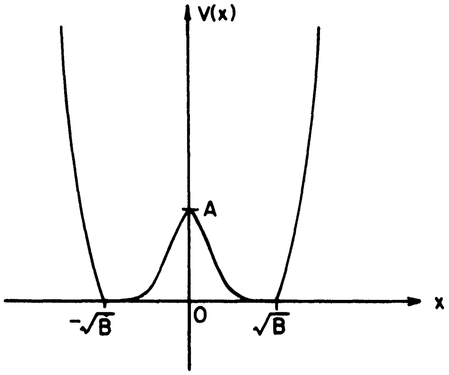


Figure 10: Sketch of the quartic potential.

$$H = \frac{p^2}{2m} + A - 2B\, x^2 + x^4 - \varepsilon x \cos \omega_o t \qquad \text{(III.1)}$$

where ε is the amplitude of the external force and ω_0 is its radial frequency. The quartic potential is plotted in Figure (10). We can introduce a canonical transformation to new variables (J,θ) which enables us to write this Hamiltonian in the form

$$H = A-B^2+J^2 - \varepsilon\sqrt{B+J}\cos(\omega_o t)\ \mathrm{dn}\left(\frac{\theta}{\sqrt{mJ}\ k'}\ \ k\right) \qquad \text{(III.2)}$$

where the modulus, k, is defined

$$k^2 = \frac{2J}{B+J} \qquad \text{(III.3)}$$

Equation (III.2) holds for all values of the modulus in the range $0<k<2$. When $\varepsilon=0$, a particle trapped in one of the wells will have energy $E<A$ and $0<J<B$, while a particle free to move between wells will have $E>A$ and $J>B$.

The perturbation term in Eq. (III.2) may be expanded in a Fourier series. For $J<B$ we have

$$H=A-B^2+J^2 - \varepsilon \sum_{n=-\infty}^{\infty} g_n(J) \cos\left[\frac{n\pi\theta}{\sqrt{mJ}\ k\ K(k)} - \omega_o t\right] \qquad \text{(III.4)}$$

where

$$g_n(J) = \frac{\pi B}{\sqrt{2}\sqrt{(2-k^2)}K}\ \mathrm{sech}\left[\frac{|n|\ \pi\ K'(k)}{K(k)}\right] \qquad \text{(III.5)}$$

and for $J>B$ we have

$$H=A-B^2-J^2-\varepsilon \sum_{n=-\infty}^{\infty} h_n(J) \cos\left[\frac{(n+1/2)\,\theta\pi}{\sqrt{mJ}\;\; K(\frac{1}{k})} - \omega_0 t\right] \tag{III.6}$$

where

$$h_n(J) = \frac{\pi\sqrt{2J}}{2kK(\frac{1}{k})} \operatorname{sech}\left[\frac{(|n| + 1/2)\pi K'(1/k)}{K\;(\frac{1}{k})}\right] \tag{III.7}$$

The locations of the primary resonance zones are determined by the condition

$$\frac{\omega_o\sqrt{m}}{2\pi n} = \frac{\sqrt{J_c}}{k_c K(k_c)} \tag{III.8}$$

for $J < B$ and

$$\frac{\omega_o\sqrt{m}}{2\pi(n+\frac{1}{2})} = \frac{\sqrt{J_c}}{K(\frac{1}{k_c})} \tag{III.9}$$

for $J>B$. A sketch of the location of the primary resonance zones as a function of frequency is given in Figure (11). We see that as $f_0 \to 0$ all primary resonance zones exist and are grouped in the neighborhood of $J \simeq B$. In the energy region $E \approx A$, there will be a stochastic layer enabling trapped particles with energies near $E=A$ to diffuse across the barrier. As frequency f_0 increases, the primary zone $n = 1$ ($J<B$) starts to drop away and move to lower energy in the phase space. At the harmonic limit of this system, it disappears. Thus, in this system we have the surprising result that as energy of a trapped particle increases the particle resonates with the external field at lower frequencies than the harmonic limit. a renormalization group analysis which quantifies these results has been performed by L.E. Reichl and W.M. Zheng.(5)

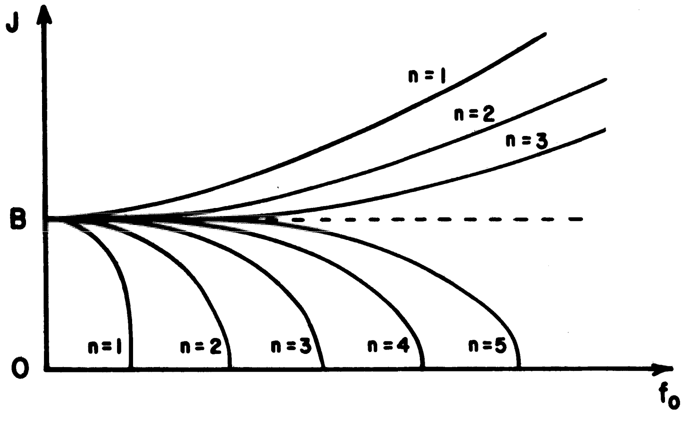


Figure 11: Sketch of J_n^c vs f_0 for n = 1,3,5,7, for the quartic potential.

IV. CONCLUSIONS

From the simple model systems we have described in previous sections, we find that dynamic external fields can profoundly distort the phase space of a nonlinear conservative system. A single frequency field induces a dense set of resonance zones throughout the phase space of the system. In practical terms this appears to have two important effects on the microscopic level. Low frequency fields (frequencies far below the natural frequencies of the unperturbed system) can lead to enhanced dissociation rates and increased rates of diffusion across barriers. External fields with frequencies in the neighborhood of the natural frequencies can in addition induce chaos and large energy fluctuations in the phase space of a nonlinear system even near the harmonic limit. This field induced chaos has the potential of disrupting coherent energy propagation (solitons) on a lattice. Perhaps the most important conceptual lesson we can learn from these models is that a nonlinear dynamical system, when perturbed by a dynamic external field, is transformed into a new system whose behavior is completely different from either of its components parts.

ACKNOWLEDGMENTS

The author wishes to thank the U.S. Air Force for its partial support of this work through Contract No. F3361S-78-D-0617.

REFERENCES

1. D.J. Kaup and A.C. Newell, Proc. R. Soc. Lond. A361, 413 (1978).
2. A partial list of references to exact solutions includes: N.G. van Kampen, J. Stat. Phys. 17, 71 (1977); L.E. Reichl, J. Chem. Phys. 77, 4199 (1982); M.O. Hongler and W.M. Zheng, J. Stat. Phys, 29, 317 (1982).
3. L.E. Reichl, R. de Fainchtein, T. Petrosky and W.M. Zheng, "Field Induced Chaos in the Toda Lattice" (Physical Review, in press).
4. R. de Fainchtein and L.E. Reichl, "Resonance Behavior of the Perturbed Toda Lattice," in Statistical Physics and Chaos in Fusion Plasmas, eds. C.W. Horton and L.E. Reichl (John Wiley and Sons, New York, in press).
5. L.E. Reichl and W.M. Zheng, "Renormalization Group Study of a Perturbed Bistable System" (Center for Statistical Mechanics, preprint).
6. J. Dancz and S. Rice, J. Chem. Phys. 67, 1418 (1977).
7. J. P. Keener and D.W. McLaughlin, Phys. Rev. A16, 777 (1977).

SOLITARY WAVES AS FIXED POINTS OF INFINITE DIMENSIONAL MAPS

David W. McLaughlin, J. V. Moloney and A. C. Newell

Department of Mathematics and
Program in Applied Mathematics
University of Arizona, Tucson, AZ 85721

INTRODUCTION

Coherent structures which are localized in space arise and play important roles in turbulent fields. Examples include tornados, large vortices in a turbulent wake, and isolated vortices in the ocean. These coherent effects are cetainly observed in nature. They can be simulated in numerical experiments and to a lesser extent in laboratory experiments. To understand such coherent phenomena analytically one must use nonlinear techniques from modern mathematics, and indeed one must extend and improve these techniques considerably.

The coexistence of coherence and chaos is a current topic in applied mathematics. Perhaps the simplist and most analytically tractable examples are solitary profiles in chaotic waves. In this lecture we will describe one example from laser optics which illustrates the very robust stability of solitary waves even as they become chaotic. Mathematically this example involves a nonlinear Schrödinger equation which is a canonical equation[1] which arises in many other contexts. For example, a form of this equation is used as a model to describe energy transport along alpha helical proteins[2,3] an application which is of direct interest to several participants of this meeting. This lecture summarizes work described in[4].

DESCRIPTION OF THE PROBLEM

Consider an optical ring cavity as is depicted in Figure

1a. An input laser signal enters a nonlinear medium at location H, travels down the medium to location I where it exits, and is reflected by four mirrors back to location H where it re-enters the nonlinear medium so as to reinforce the original input signal. The signal continues to circulate around the cavity; our goal is to predict the output after many passes through the cavity.

Mathematically the classical electromagnetic field is described by Maxwell's equations while the polarization of the medium is described quantum mechanically by a form of the Schrödinger equation known as the Bloch equations. Since the electromagnetic field is a rapidly oscillating plane wave with a slowly varying envelope G,

$$\vec{E}(x,z,t) = 2\,\mathrm{Re}[G(x,z,t)e^{i(kz-\omega t)}]\hat{x}, \qquad (1)$$

the coupled system which describes the interaction of the electromagnetic field with the medium may be approximated by a single nonlinear partial differential equation for the envelope G. In scaled variables this equation takes the form

$$2i\left(\frac{\partial}{\partial x} + \frac{1}{c}\frac{\partial}{\partial t}\right)G + \frac{\partial^2}{\partial x^2}G - \frac{G}{1+2|G|^2} = 0\ . \qquad (2)$$

Equation (2) describes the propagation of the electric field through the nonlinear medium. It must be supplemented by a boundary condition which initializes the field at entry point H. This boundary condition will depend upon the input laser signal and upon the field G that has just emerged from the medium at the preceding pass. This emerging field is then modified from its value at exit point I by losses at the mirrors and by the optical path length of the return trip from I to entry point H. Using the fact that signals propagate (in z) at speed c, the field on the nth pass through the medium satisfies

$$\begin{aligned} &2i\frac{\partial}{\partial\tau}G_n + \frac{\partial^2}{\partial x^2}G_n - \frac{G_n}{1+2|G_n|^2} = 0 \\ &G_n(x,0) = a(x) + Re^{ikL}G_{n-1}(x,P), \end{aligned} \qquad (3\ a,b)$$

$n > 1$, $G_0 = 0$. Here we have transformed to a moving frame (z-ct, τ=ct), a(x) denotes the envelope of the input laser signal, P measures the length of the nonlinear medium, L measures the optical path length of the return trip from I to H, and $R \simeq .9$ denotes losses at the mirrors.

Equation (3) defines the mathematical problem. At this point it is worthwhile to pause and make several remarks: (i)

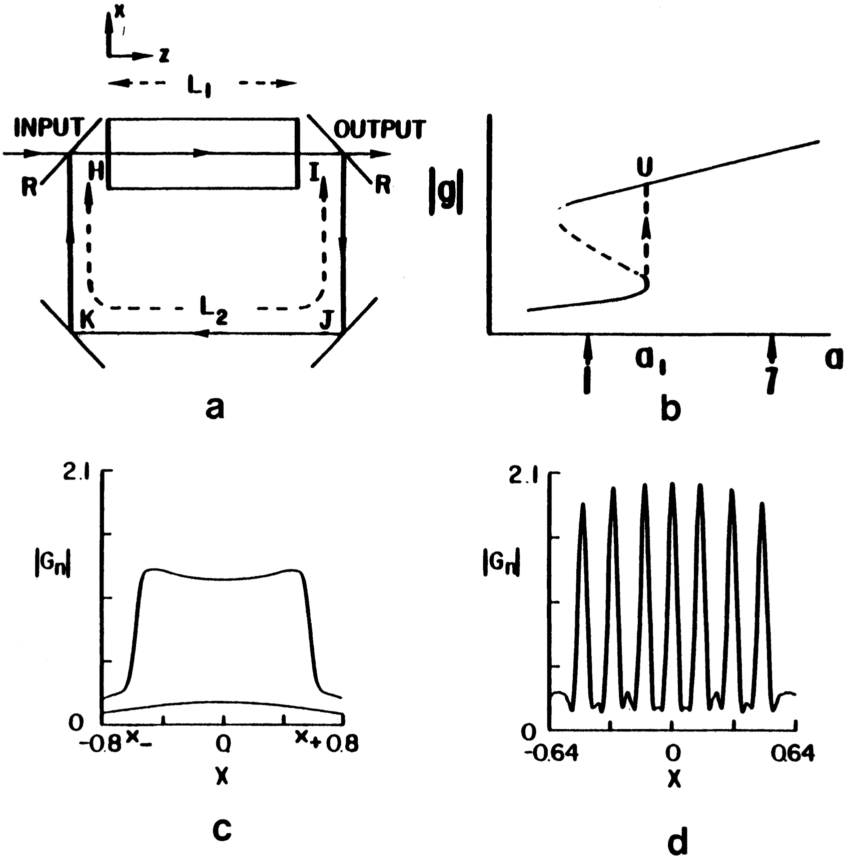


Figure 1

Figure 1. (a) Unidirectional passive ring cavity containing a nonlinear medium. (b) Plane wave bistable loop. (c) Initiation of the sharp gradients on the outer edges of the transverse laser beam. (d) Steady state 7-solitary wavetrain at a(0) = .194.

Problem (3) defines an infinite dimensional map from the field at the end of the $(n-1)^{st}$ pass to the field at the end of the n^{th} pass,

$$G_{n-1}(x,P) \rightarrow G_n(x,P). \qquad (4)$$

(ii) This map is infinite dimensional because of its transverse x dependence. In order to find G_n in terms of G_{n-1}, one must solve the partial differential equation (3a). (iii) In order to study analytically the behavior of G_n for large n, we must learn to reduce the dimension of map (4) in a manner which is natural to problem (3). Solitary wave solutions of equation (3a) will turn out to be the natural objects with which to carry out this reduction of dimension.

PLANE WAVE CASE

Another way to reduce the dimension of map (4) is to assume that the input laser beam has a transverse profile a(x) which is constant, independent of x. This situation is known as the "plane wave" case. Although it is not a particularly natural reduction, it is a necessary prerequiste for that which follows.

If the input "a" is independent of x, so is the output G_n. In this case equation (3a) reduces to an ordinary differential equation which can be solved in closed form, and map (4) can be written explicitly:

$$g_{n+1} = a + R \exp[i(kL - \frac{P}{2(1+2|g_n|)^2)})]\, g_n. \qquad (5)$$

This is a one dimensional complex (or a two dimensional real) map. It has stable fixed points which represent a balance between the energy injected by the input signal "a" and the energy lost at the mirrors. Multiple fixed points can arise which are the result of a resonance which allows the returned beam to interfere constructively with the input signal. For a range of a and P, map (5) possesses exactly three fixed points (Figure 1.b), and it exhibits a bistable behavior. Since the lower and upper branch are stable, one has the possibility of an optical switch. (Optical scientists[5] are interested in such devices because these switches are potentially much faster than conventional ones.)

Map (5) possesses a wide variety of behavior which is much richer than stable and unstable fixed points. For example, period doubling sequences[6] leading to chaotic behavior have been observed[7]. At present the universal properties of map (5) have not been analytically described.

To close this discussion of the plane wave map (5), we return to Figure 1.b. As the input amplitude "a" is increased past a_1, the output field jumps to the upper branch at U. Now consider an input field whose transverse profile depends upon x, a = a(x). If this dependence is gradual (as, for example, a broad Gaussian profile), the Laplacian ∂_{xx} in (3a) is initially negligible. At each x the beam behaves as if it were a uniform plane wave at input amplitude a(x). Thus, the center of the transverse profile (for which $a(x) > a_l$) will switch up to the upper branch, while the edges (for which $a(x) < a_l$) will stay on the lower branch. Therefore, at $x_\pm$ where $a(x_\pm) = a$, the gradient of the response field is very large (see Figure 1c). Near these locations the Laplacian ∂_{xx} cannot be neglected and the plane wave theory becomes inapplicable.

What actually happens? Numerical experiments show that during the first twenty or so passes $(0 \leqslant n \leqslant 20)$, the Gaussian profile switches to the square profile of Figure 1.c as predicted by the plane wave theory. Then at the edges x_+ and x_- narrow pulses are generated which eventually fill out the area between x_- and x_+ and become the steady state response of the system, Figure 1.d. The number of pulses in the steady state response depends upon parameter values. We have observed between 1 and 7.

SOLITARY WAVE REDUCTION

We emphasize that these pulses or rings in the transverse profile are fixed points of the infinite dimensional map (4). This pulse structure is very stable. Consider the output of the n^{th} pass at exit position I. Its transverse profile is a collection of pulses. This shape is damaged as it returns from I to H where it will re-enter the nonlinear medium. This damage consists of (i) decreasing its amplituide by the factor R due to losses at the mirrors, (ii) shifting its phase by the optical path length, and (iii) adding the small input field a(x). Although this distorted field enters the nonlinear medium at the start of the $(n+1)^{st}$ pass, by the time that it emerges from the nonlinear medium at I, it is exactly the same as it was at the end of the n^{th} pass.

We will argue that the shapes of the pulses in the transverse profile are fixed by equation (3a). These shapes are solitary wave solutions of equation (3a). However, since equation (3a) is conservative, it cannot select the amplitudes of these rings. The amplitude of that ring which describes a fixed point of (4) will be selected by the map nature of boundary condition (3b).

Consider a nonlinear Schrodinger equation of the form

$$2i \frac{\partial}{\partial \tau} G + \frac{\partial^2}{\partial x^2} G + N(|G|^2)G = 0. \qquad (6)$$

Equation (3a) is of this form with the nonlinearity $N(|G|^2) = -(1+2|G|^2)^{-1}$. A solitary wave is a one parameter family of solutions of (6) in the form

$$G(x,\tau;\lambda,\gamma) = S(x;\lambda)\exp[i(\frac{\lambda^2-1}{2} + \gamma)], \qquad (7)$$

where $S(\theta;\lambda)$ is a solution of

$$S'' - S + \frac{N(S^2)+1}{\lambda^2} S = 0, \qquad (8)$$

which vanishes (exponentally) as $\theta \to \pm\infty$. The simplest example is that of a cubic nonlinearity $N(S^2) = -1 + S^2$, in which case

$$S(\theta;\lambda) = \lambda \operatorname{sech}\theta \quad \text{(cubic nonlinearity).} \qquad (9)$$

From this we see that the parameter λ measures the amplitude of the solitary wave. For our nonlinearity $N(S^2) = -(1 + 2S^2)^{-1}$ no analytical formula such as (9) exists; however, $S(\theta;\lambda)$ still has a shape which is qualitatively similar to a sech and with the parameter λ fixing its amplitude. This, we see that the solitary wave (7) is a wave form which is localized in x and which depends upon two parameters λ and γ. λ fixes the amplitude (and width) of the pulse; γ fixes its phase.

If we assume that the output of the nth pass through the medium is a solitary wave with parameters (λ_n, γ_n), then the input for the (n+1)st pass is given by (3b),

$$G_{n+1}(x,0) = a(x) + RS(\lambda_n x;\lambda_n)e^{i[\gamma_n + kL + \frac{P}{2}(\lambda_n^2-1)]}, \qquad (10)$$

which may be rewritten as

$$G_{n+1}(x,0) = S(\lambda_n x;\lambda_n)e^{i[\gamma_n + kL + \frac{P}{2}(\lambda_n^2-1)]} + \tilde{G}_{n+1}(x) \qquad (10')$$

where

$$\tilde{G}_{n+1}(x) = (R-1)S(\lambda_n x;\lambda_n)e^{i[\gamma_n + kL + \frac{P}{2}(\lambda_n^2-1)]} + a(x) \qquad (11)$$

Since both (R-1) and a(x) are small, $\tilde{G}_{n+1}$ is small. Thus, initial data (10') for the (n+1)st pass is a small perturbation of a solitary wave with parameters $(\lambda = \lambda_n, \gamma = \gamma_n + kL + \frac{P}{2}(\lambda_n^2-1))$. Using a projection method from solitary wave perturbation theory, one can predict to first order the parameters $(\lambda_{n+1}, \gamma_{n+1})$ of the solitary wave which emerges at the end of the (n+1)st pass; that is, one can compute a map

$$(\lambda_n,\gamma_n) \to (\lambda_{n+1}, \gamma_{n+1}). \qquad (12)$$

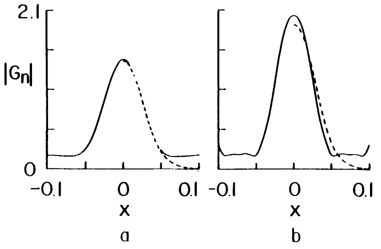


Figure 2. (a) Comparison of the single solitary wave shape (N=1) at a(0) = .1 predicted from the fixed point equation (12) (dashed line) with the numerical solution of equation (3) (over 200 cavity passes) (solid line). (b) The central peak of the 7-solitary wavetrain (see Figure 1d) is compared to the shape predicted by the fixed point equation (a(0) = .194). The discrepancy in fit ($\sim$ 5%) is consistent with perturbative estimates of changes in shape due interactions with nearby solitary waves.

The actual formulas for this map are somewhat complicated and will not be reproduced here. Rather we restrict ourselves to a several comments about its properties: (i) Map (12) is a two dimensional real map. Using solitary waves which are the natural nonlinear modes of equation (3a) we have reduced the infinite dimensional map (4) to a two dimensional one. (ii) Although it is more complicated, the algebraic structure of map (12) is very similar to that of the plane wave map (5) which is also a two dimensional real map. (iii) Map (4) has a fixed point (λ_*,γ_*) which we found numerically. This fixed point predicts the amplitude of the central ring which emerges after many passes through the nonlinear medium. In Figure 2 we compare this prediction of map (12) with the numerical solution of (3). The prediction is extremely accurate!

Finally, we describe recent results not contained in ref. 4.

Recall that the plane wave map (5) displayed a period doubling route to deterministic chaos. Because map (12) is similar in structure to map (5), we sought period doubling this ringed case numerically. We concentrated on a parameter range for which the fixed point of the infinite dimensional map (4) is a single solitary wave. We found that as the input amplitude parameter a(0) is increased, the single solitary wave fixed point becomes unstable and a period 2 (in n) state describes the large n asymptotic behavior. The transverse profile of this period 2 state remains a single solitary wave. As the parameter a(0) continuous to increase, we observe numerically an "almost" period 4 state and chaotic states. The transverse profiles remain predominately solitary waves although they do possess a little additional structure.

CONCLUSION

For the purposes of this conference, it is important to emphasize the remarkable stability of the solitary wave as is illustrated in this example from laser physics. Even in the presence of deterministic chaos, the profile remains a coherent, localized solitary wave. This coherent nonlinear pulse dominates the qualitative behavior of the large n output. Judicious use of solitary waves permits a theoretical description of features which linear methods would not capture.

REFERENCES

1 David W. McLaughlin, Solitary waves with dispersion and dissipation in: "Structure and Dynamics: Nucleic Acids and Proteins," E. Clementi and R.H. Sarma, ed., Adenine Press, New York (1983).

2. A.S. Davydov, Energy transfer along alpha-helical proteins in: "Structure and Dynamics: Nucleic Acids and Proteins," E. Clementi and R. H. Sarma, ed., Adenine Press, New York (1983).

3. A.C. Scott, Solitons on the alpha-helix protein in: "Structure and Dynamics: Nucleic Acids and Proteins," E. Clementi and R. H. Sarma, ed., Adenine Press, New York (1983).

4. D.W. McLaughlin, J.V. Moloney, and A.C. Newell, Solitary waves as fixed points of infinite-dimensional maps in an optical bistable ring cavity, Phys. Rev. Lett. 51: 75 (1983).

5. H. Gibbs, Optical logic--what is it? where does it fit in?, Preprint, Optical Sciences Center, University of Arizona (1982).

6. M. Feigenbaumm, The transition to aperiodic behavior in turbulent systems, Comm. Math. Phys. 77:65 (1980).

7. K. Ikeda, Mutiple-valued stationary state and its instability of the transmitted light by a ring cavity system, Opt. Comm. 30: 257(1979).

NONLINEAR PHENOMENA IN A DISPERSIVE SYSTEM WITH APPLICATIONS TO WATER WAVES

Henry C. Yuen

Fluid Mechanics Department
TRW Space and Technology Group
Redondo Beach, California 90278

ABSTRACT

Significant progress in the understanding of nonlinear dynamics has been made in the past decade. New concepts, such as solitons, recurrence, chaotic motion and bifurcation have been discovered and related to important physical phenomena. The study of nonlinear deep-water waves has been both the cause and consequence of many of these novel concepts.

A survey of the recent developments in water waves is presented. This includes the nonlinear Schrödinger equation, the theoretical and experimental studies of envelope solitons and recurrence phenomena, stability analysis in two and three dimensions, discovery of new three-dimensional steady wave patterns and chaotic behavior in the long time evolution of a nonlinear wavetrain. Much of the mathematics and many of the techniques are related and applicable to other branches of science, including biological systems.

INTRODUCTION

One of the most vivid examples of a nonlinear dispersive wave system is deep water waves. New concepts, such as solitons, recurrence and chaotic behavior are all present in the description of wave packets and wavetrains on deep water. Compared with many other physical, chemical or biological systems, water waves possess the unique advantage of being familiar and easily measurable. In this paper, we present highlights of some recent progress in the study of water waves as illustrations of the type of behavior encountered in a nonlinear dispersive wave system.

THE GOVERNING EQUATIONS

The equations governing irrotational flow of an incompressible inviscid fluid with a free surface are:

$$\Delta\phi = 0 \qquad -\infty < z < \eta(x,y,t) \tag{1}$$

$$\left.\begin{aligned} &\frac{\partial\phi}{\partial t} + \tfrac{1}{2}(\nabla\phi)^2 + gz = p \\ &\frac{\partial\eta}{\partial t} + \nabla\phi\cdot\nabla\eta - \frac{\partial\phi}{\partial z} = 0 \end{aligned}\right\} \qquad z = \eta(x,y,t) \tag{2}$$

$$\frac{\partial\phi}{\partial z} \to 0 \qquad z \to -\infty \tag{3}$$

where ϕ is the velocity potential, η is the free surface, g is the gravitational acceleration, p is an external surface pressure, the horizontal coordinates are (x,y) and the vertical coordinate is z, pointing upwards; $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is the Laplacian operator, and $\nabla = (\partial/\partial x, \partial/\partial y)$ is the horizontal gradient operator. The density is normalized to unity with no loss of generality. For a wave to be considered a deep-water wave, the ratio of its wavelength λ to the water depth must be small. For surface tension to be negligible (i.e., for a wave to be considered a gravity wave), the wavelength must be substantially longer than 1.7 cm, at which point surface tension and gravitational effects are equal. Strictl speaking, water is not inviscid. However, viscosity is most effective only for small-scale motions and is negligible for most of the gravity-wave phenomena considered here. Therefore, potential flow provides a good approximation.

The fundamental solutions for the linearized problem are obtained by linearizing the boundary conditions at the undisturbed surface z = 0,

$$\eta(x,t) = a\cos(\underset{\sim}{k}\cdot\underset{\sim}{x} - \omega t) \tag{4}$$

$$\phi = \frac{ga}{\omega}\sin(\underset{\sim}{k}\cdot\underset{\sim}{x} - \omega t)e^{|\underset{\sim}{k}|z}, \tag{5}$$

where $\underset{\sim}{x} = (x,y)$ is the horizontal spatial vector, $\underset{\sim}{k} = (k_x,k_y)$ is the wave vector, and ω the wave frequency in radians. The linear dispersion relation, which relates the frequency to the wave vector, is given as

$$\omega = \sqrt{g|\underset{\sim}{k}|}. \tag{6}$$

This defines the phase velocity, which gives the speed of advance of the individual crests, as

$$C = \frac{\omega}{\underset{\sim}{k}} = \frac{\omega}{|k|^2}(k_x,k_y) \tag{7}$$

and the group velocity, which describes the speed and direction of energy propagation, as

$$C_g = \frac{\partial\omega}{\partial\underset{\sim}{k}}\left(\frac{\partial\omega}{\partial k_x}, \frac{\partial\omega}{\partial k_y}\right). \tag{8}$$

The fact that C is a nontrivial function of $\underset{\sim}{k}$ reflects the dispersive nature of the system.

Henceforth, unless stated otherwise, we shall be discussing one-dimensional waves that are independent of y. For simplicity of notation, we shall denote k_x by k in these one-dimensional problems.

CONCEPT OF A WAVETRAIN

A wavetrain can be characterized by three parameters: its amplitude a, wavenumber k, and frequency ω. Generally, its evolution is described by equations for the conservation of wavenumber, the conservation of wave action, and the dispersion relation. For weakly nonlinear waves, they are respectively

$$\frac{\partial k}{\partial t} + \frac{\partial\omega}{\partial k} = 0 \tag{9}$$

$$\frac{\partial a^2}{\partial t} + \frac{\partial}{\partial x}\left(\frac{\partial\omega}{\partial k}a^2\right) = 0 \tag{10}$$

$$\omega = \sqrt{gk}\,(1 + \tfrac{1}{2}k^2a^2). \tag{11}$$

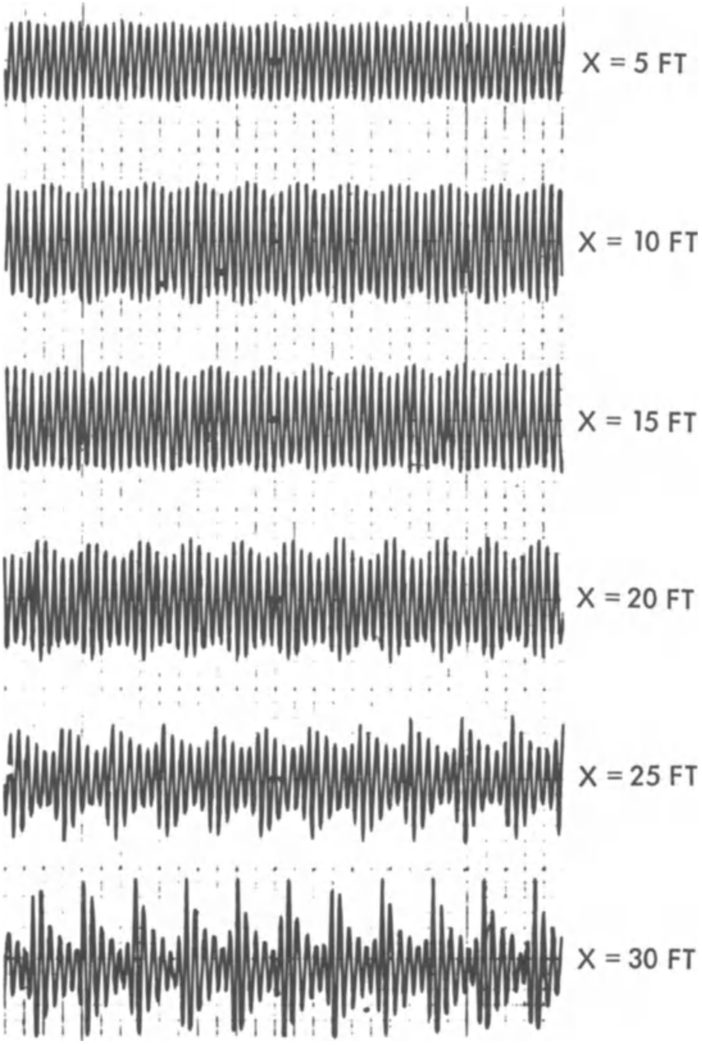


Figure 1. Wave amplitude measurements showing the onset of the Benjamin-Feir modulational instability on an initially uniform wavetrain. The distance x is measured from the wave paddle. The gauges do not have the same calibration and hence relative differences in magnitudes for measurements at different x locations do not infer changes in wave amplitudes. (TRW Fluid Mechanics experiment.)

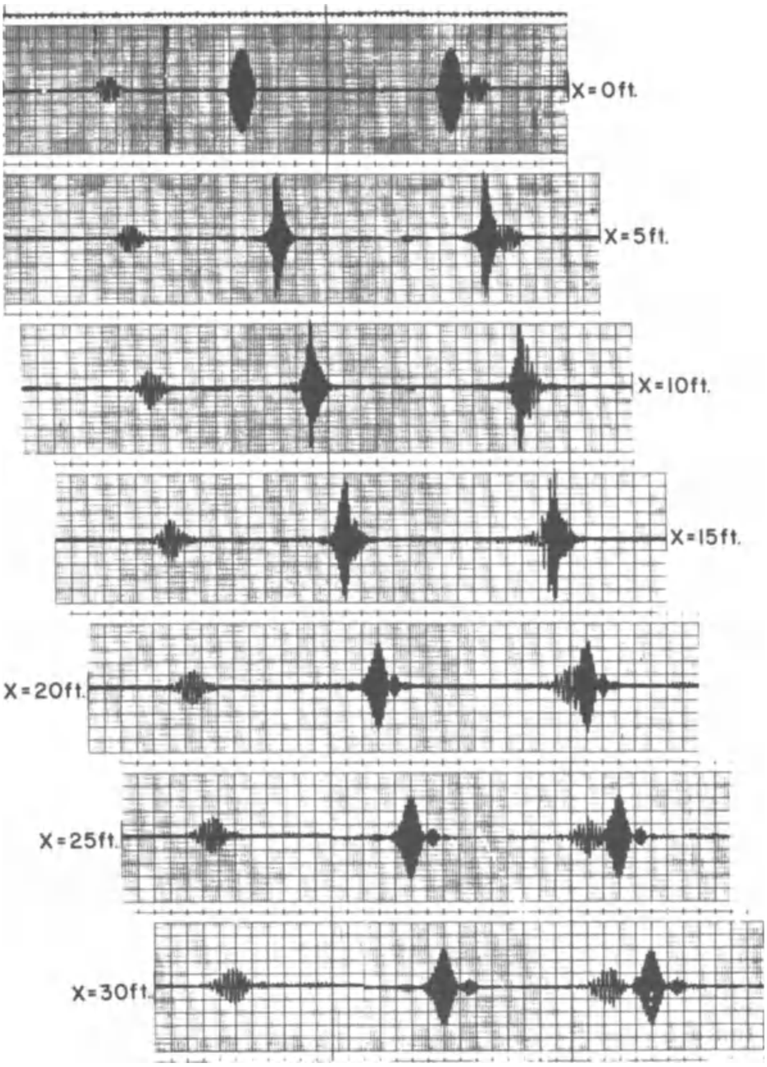


Figure 2. Example of interaction of two wave packets. Left-hand trace: packet with $\omega_o = 1.6$ Hz, $k_oa_o = 0.1$, propagating alone; Center trace: packet with $\omega_o = 3$ Hz, $k_oa_o = 0.2$, propagating alone and disintegrating into more than one "soliton"; Right-hand trace: 3 Hz packet followed immediately by the 1.6 Hz packet; the faster (due to lower frequency) packet overtakes and interacts with the slower packet, but they both emerge without permanent change of form.

These are the leading-order equations derivable from the averaged variational principle of Whitham (1965, 1967, 1970) and are sometimes known as Whitham's conservation equations. Equivalent systems can be derived by a variety of methods, including the multiple-scales method (Chu & Mei 1971).

Note that the dispersion relation contains a nonlinear correction that was first obtained by Stokes (1849). The absolute magnitude of the correction is small, and does not exceed 11% even for the steepest possible waves. Nevertheless, its existence couples the three equations nonlinearly, and leads to some remarkable phenomena during the evolution of the weakly nonlinear wavetrain.

MODULATIONAL INSTABILITY

Equations (9)-(11) were used by Lighthill (1965) to study the evolution of weakly nonlinear waves on deep water. Two sets of initial conditions were examined: a wave packet with a Gaussian envelope and a slightly modulated Stokes wavetrain. In the case of the wave packet, the envelope developed a cusp at its peak within a finite time. For the wavetrain, the modulation grew and again the envelope developed into a cusped shape. Smooth solutions ceased to exist beyond this stage of evolution, and the time taken to reach the singularities was found to be inversely proportional to the steepness of the waves.

These results were the first indications of modulational instabilities for weakly nonlinear waves on deep water. A detailed perturbation analysis of the uniform wavetrain performed on the Euler equations (1)-(3) was presented by Benjamin & Feir (1967) that confirmed the instability and extended the results to a wider range of perturbation wavelengths. It was found that a uniform wavetrain with amplitude a_o, wavenumber k_o, and frequency ω_o is unstable to perturbations with wavenumber K in the range

$$0 < K < 2\sqrt{2}\, k_o^2 a_o . \tag{12}$$

The maximum instability occurs when $K = K_{max}$, where

$$K_{max} = 2k_o^2 a_o , \tag{13}$$

and the maximum growth rate $(\mathrm{Im}\Omega)_{max}$ is

$$(\mathrm{Im}\Omega)_{max} = \tfrac{1}{2}\, \omega_o k_o^2 a_o^2 . \tag{14}$$

Benjamin & Feir (1967) also reported experimental data in fairly good agreement with the predictions regarding the wavenumber and growth rate of the instability. An example of such a modulation growing on an initially uniform wavetrain is shown in Figure 1.

THE NONLINEAR SCHRÖDINGER EQUATION

Identification of the instability led to the question of the subsequent evolution of the unstable nonlinear wavetrain, since the results of Benjamin & Feir were confined to the initial growth period. Hasimoto & Ono (1972) used the derivative-expansion method to show that the evolution of a nonlinear deep-water wavetrain obeys the nonlinear Schrödinger equation, and they recovered the results of Benjamin & Feir by a stability analysis of its uniform solution. Yuen & Lake (1975) confirmed that Whitham's equations, when the dispersion relation is included, lead to the same nonlinear Schrödinger equation for the complex wavetrain envelope $A(x,t) = a(x,t)e^{i\theta(x,t)}$:

$$i\left(\frac{\partial A}{\partial t} + \frac{\omega_o}{2k_o}\frac{\partial A}{\partial x}\right) - \frac{\omega_o}{8k_o^2}\frac{\partial^2 A}{\partial x^2} - \tfrac{1}{2}\,\omega_o k_o^2|A|^2A = 0 \qquad (15)$$

where $\partial\theta/\partial t = \omega_o - \omega$, $\partial\theta/\partial x = k - k_o$, and the free surface of water is $\eta(x,t) = \mathrm{Re}\{A \exp i(k_o x - \omega_o t)\}$.

ENVELOPE SOLITONS

One of the most important characteristics of the nonlinear Schrödinger equation is that it can be solved exactly for initial conditions that decay sufficiently rapidly as $|x|\to\infty$. This was done by Zakharov & Shabat (1971) using what was then the newly discovered inverse scattering method (Gardner et al 1967). They showed that any initial wave packet eventually evolves into a number of "envelope solitons" and a dispersive "tail". The bulk of the energy is contained in the solitons, which propagate with permanent form once produced. Solitons also survive interactions with other solitons or wave packets. Since the nonlinear Schrödinger equation describes the envelope of deep-water waves with a carrier frequency, the theory predicts the existence of packets of deep-water waves with soliton properties.

The theoretical predictions of soliton properties were confirmed by experiments (an example is shown in Figure 2), and numerical solutions of the nonlinear Schrödinger equation were found to compare well with experimentally measured wave-envelope profiles (Yuen & Lake 1975). In the case of wave packets and solitons, the description of nonlinear deep-water wave dynamics provided by the nonlinear Schrödinger equation was found to be both qualitatively

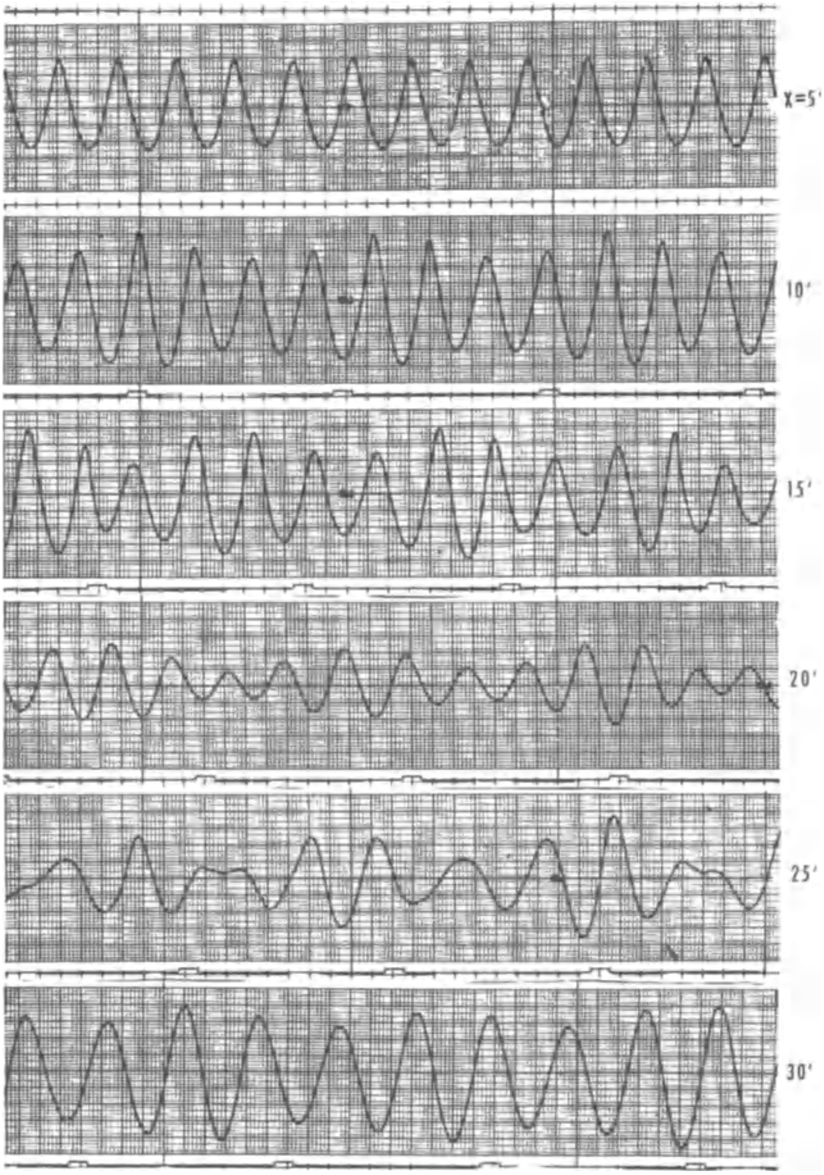


Figure 3. Evolution of an initially uniform wavetrain through one cycle of modulation and demodulation. (TRW Fluid Mechanics Department experiment)

and quantitatively correct. Although the practical significance of these envelope solitons remains to be determined because of their unidirectional nature, their mathematical and physical existence increased confidence in the nonlinear Schrödinger equation description and provided a dramatic demonstration of the potential significance of nonlinear effects on the dynamics of deep-water gravity waves.

LONG-TIME EVOLUTION OF AN UNSTABLE WAVETRAIN

Experiments in the laboratory by Lake et al (1977) indicated that the unstable modulations grew to a maximum and then subsided (Figure 3). Furthermore, at some stage of the evolution the wavetrain actually became nearly uniform again. Numerical computations using the nonlinear Schrödinger equation, which proved to be satisfactory for describing the long-time evolution of the wave packets, confirmed this interesting phenomenon: In the absence of viscous dissipation there are no steady end-states, but a series of modulation and demodulation cycles, known as the Fermi-Pasta-Ulam recurrence phenomenon (Fermi, Pasta & Ulam 1955; aee also Scott, Chu & McLaughlin 1973).

A positive link between the character of the long-time evolution and the initial conditions was reported by Yuen & Ferguson (1978a). They showed by numerical experiment that the long-time evolution of an unstable wavetrain is dominated by the unstable modes contained in the initial condition. The stable modes do receive some energy, but they never appear as the dominant mode at any stage of evolution.

To illustrate this relationship, we present several cases in which different numbers of harmonics fall within the unstable regime. We use the notation that case n corresponds to a case in which n modes, including the fundamental, are unstable and thus expected to actively participate in the evolutionary process. The values of K for each case ($1 \leq n \leq 5$) are given in Figure 4(a) and the results of the computation are shown in Figure 4(b) as plots of the envelope versus space and time. Note that the number of modulational peaks present in the evolution of a case corresponds to the number of unstable modes n.

CHAOS AND ORDER IN LONG TIME EVOLUTION

In the case where the initial condition contains no unstable modes, the wavetrain is stable, and the envelope remains as a constant. In the case where the initial condition contains one unstable mode, a recurrence sequence is set up. The regularity of the recurring sequence depends on the detail of the initial conditions and the number of stable modes involved, but the dependence is not sensitive. In the case where the initial condition contains more than one unstable mode, the evolution process is highly irregular, and the evolution history is extremely sensitive not only to the number of unstable modes involved (which is expected), but also to the number of stable modes included. The latter dependence on stable modes is somewhat surprising because these stable modes contain insignificant energy throughout the evolution, yet their mere presence appears to strongly affect the qualitative structure of the evolutionary process. The system can be considered "chaotic", not in the popular sense of sensitive dependence on initial condition, but in the sense that the evolution is sensitive to "small perturbations" in the form of stable modes receiving only minimal energy throughout. The nature of the evolution of an inviscid wavetrain is still under investigation.

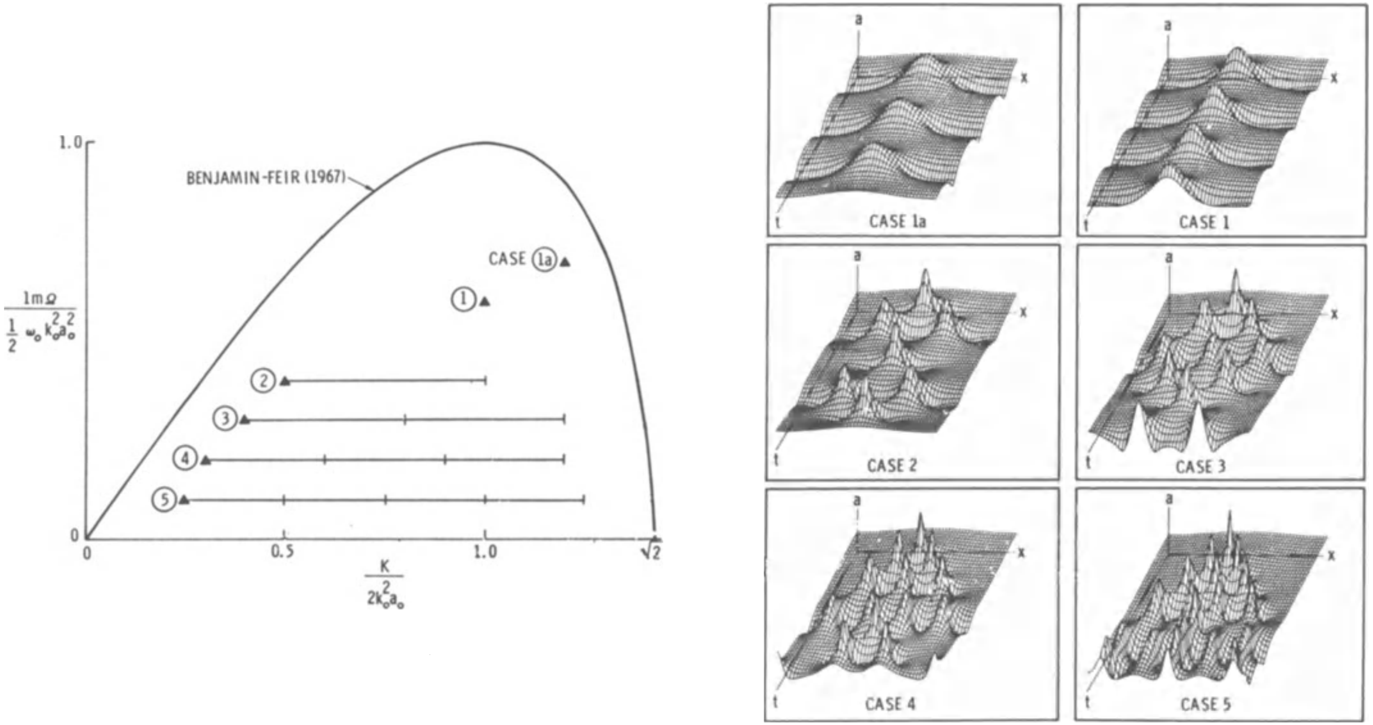
The manner in which the evolution depends on the number of stable modes is also surprising. It was found that the increase in the number of stable modes, even though each pair brings with it new and higher frequencies, appears to reduce the degree of irregularity of the time evolution. In Figure 5, we show the plots of the real and imaginary parts of the dominant modes (left hand traces) and the absolute values of the dominant versus the most unstable modes (right hand traces) of three cases. The first is a two-unstable-mode chaotic case, where the dominant wave is perturbed by noise level disturbances of 10 other modes, only 2 pairs of which are unstable, all of magnitude 10^{-6} compared with the dominant. The second is the same case, except that a total of 30 perturbation modes are included (again only 2 pairs of which are unstable). The evolution is far less irregular, and in fact, resembles the third case, in which only one pair of unstable modes is present, and is therefore recurring rather than disorderly. These results are interesting and are still presently under investigation.

With the introduction of viscosity, it has been found that the sensitivity to the inclusion of stable modes is reduced. One explanation is that viscosity effectively removes the influence of these stable modes on the evolutionary process.

WAVETRAIN IN TWO SPACE DIMENSIONS

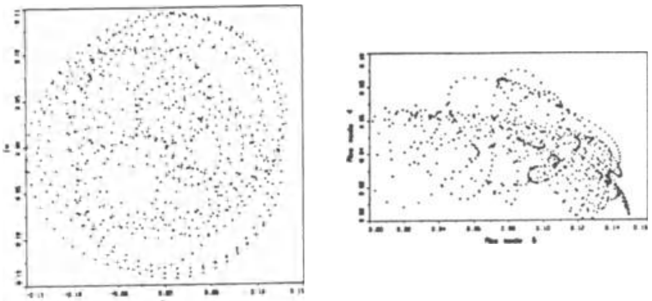
Thus far we have confined our discussion to one-dimensional problems. The equation for the complex envelope of a wavetrain propagating in two space dimensions is

$$i\left(\frac{\partial A}{\partial t} + \frac{\omega_o}{2k_o}\frac{\partial A}{\partial x}\right) - \frac{\omega_o}{8k_o^2}\frac{\partial^2 A}{\partial x^2} + \frac{\omega_o}{4k_o^2}\frac{\partial^2 A}{\partial y^2} - \tfrac{1}{2}\,\omega_o k_o^2 |A|^2 A = 0 \qquad (16)$$

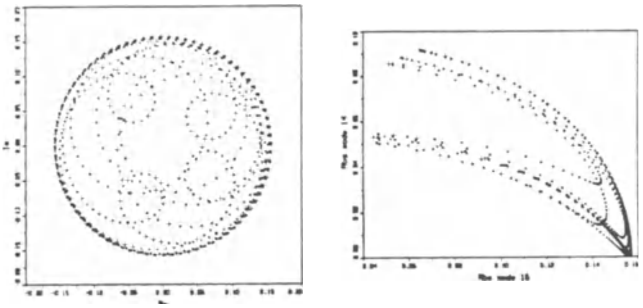


a **b**

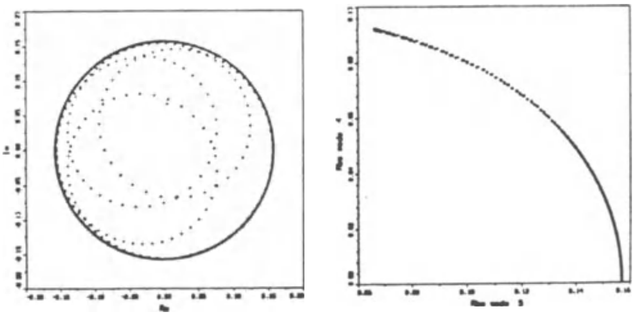
Figure 4. Relationship between initial condition and long-time evolution of the nonlinear Schrödinger equation. Figure 4a shows the location of the perturbation wavenumber K in the initial conditions for cases 1 through 4 as marked by ▲. Figure 4b shows the evolution for each of the cases. Note that the number of modulational peaks appearing for a given case is equal to the total number of unstable harmonics present in the initial perturbation as given by the case numbers.



DISORDER (unstable modes: 2; total modes: 11)



RETURN TO ORDER? (unstable modes: 2; total modes: 31)



RECURRING (unstable modes: 1; total modes: 9)

Figure 5. Order and disorder in the time evolution of a case with one and two unstable modes. Left-hand traces are plots of real versus imaginary parts of the dominant modes, right-hand traces are plots of the absolute values of the dominant mode against that of the most unstable mode. Top: case with two unstable modes and 11 total modes; Middle: case with the same two unstable modes but with 31 total modes; Bottom: case with only one unstable mode. The bottom case (one unstable mode) corresponds to recurrence and is orderly. However, it is apparent that the increase in the total number of modes appears to reduce the level of disorder.

where the free surface of water is

$$\eta(x,y,t) = \text{Re}\left\{A(x,y,t)e^{i(k_o x-\omega_o t)}\right\} . \tag{17}$$

When there is no y-dependence in the wave envelope, the equation reduces to the nonlinear Schrödinger equation (15). Note that the second-y-derivative term and the second-x-derivative term are of opposite sign, leading to a basically hyperbolic, rather than an elliptic, equation. Also note that the carrier wave still propagates in the x-direction, but now its envelope entertains two-dimensional variations.

In the following we briefly review the concepts of envelope solitons, modulational instability, and recurrence in the two-dimensional context.

Envelope Solitons Plane envelope solitons exist for Equation (16) provided that the angle between the directions of propagation of the carrier wave and of the envelope profile is less than 35.26°. An illustration of an oblique plane soliton is given in Figure 6. These two-dimensional plane solitons are infinite in extent. In addition, they are unstable to two-dimensional disturbances (Zakharov & Rubenchik 1973, Saffman & Yuen 1978). Thus far, no fully two-dimensional solutions exhibiting true soliton behavior have been found.

Modulational Instability Generalization of the stability analysis of a uniform wavetrain to two space dimensions yields the result illustrated in Figure 7. The Benjamin-Feir results are recovered by setting K_y to zero. Of particular significance is the fact that the two-dimensional instability does not have a high-wavenumber cutoff. This result has important implications when the long-time evolution of an unstable wavetrain in two dimensions is considered.

Long-Time Evolution and Recurrence For certain initial conditions, the phenomenon recurrence apparently still exists for wavetrains in two dimensions (Yuen & Ferguson 1978b, Figure 8). When one examines the relationship between initial conditions and long-time evolution one concludes that, unlike the one-dimensional result, there exists for any fully two-dimensional unstable perturbation vector (K_x,K_y) an infinite set of integers (m,n) such that (mK_x,nK_y) lie in the unstable regime and are eligible to participate actively in the evolution. Furthermore, it has been confirmed by Martin & Yuen (1980) that an unstable wavetrain does recur, but that during each new cycle, unstable modes with larger values of m and n are excited. Thus, the long-time evolution consists of a periodic return to the initial condition together with a gradual "leak" of energy to higher

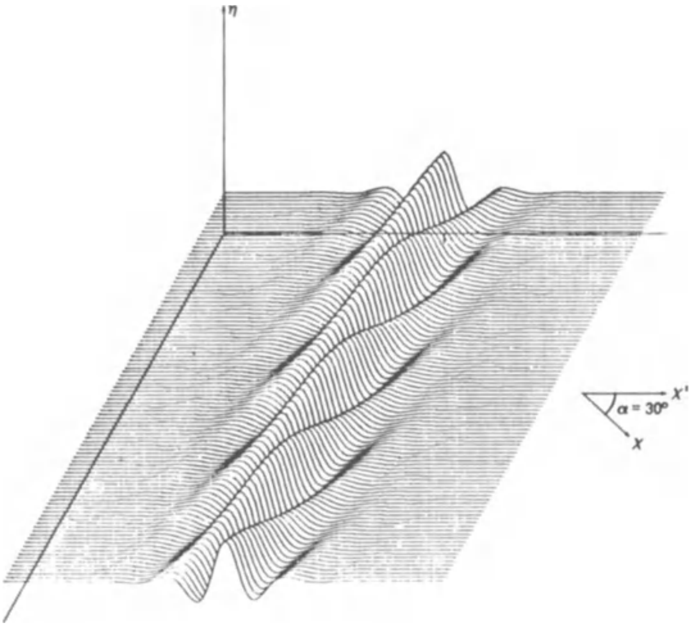


Figure 6. Sketch of the free surface associated with an oblique plane soliton with individual wave crests having an angle of 30° with the envelope.

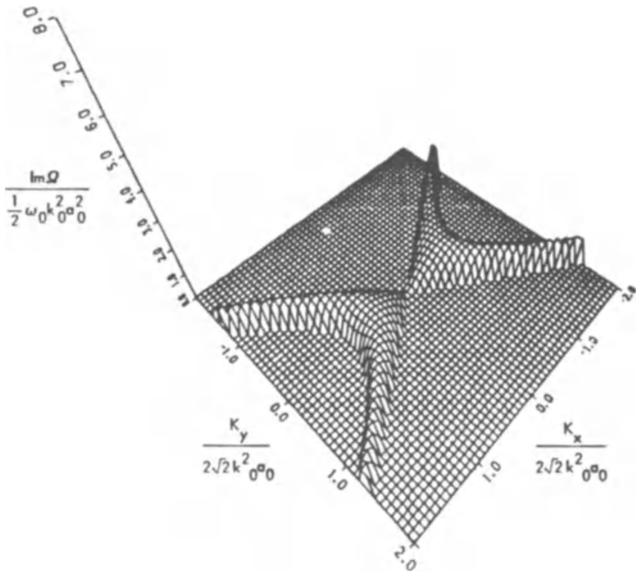


Figure 7. Stability diagram of a uniform wavetrain subject to infinitesimal two-space dimensional perturbations. The growth rate is measured by ImΩ.

and higher modes. Although this may not be the classical definition of thermalization, such behavior clearly indicates that the validity of the two-dimensional equation ceases at a finite time, since it is derived on the premise that most of the energy is confined within a narrow band around the carrier wavenumber.

ZAKHAROV'S INTEGRAL EQUATION

Most of the desirable features of the solutions of the one-dimensional nonlinear Schrödinger equation are lost in the extension to two space dimensions. The inverse-scattering method can no longer be applied. No localized solitons exhibiting stability and immunity from interaction effects have been found. Most perturbing of all, however, is the fact that certain solutions of the two-dimensional nonlinear Schrödinger equation experience a progressive leakage of energy to high modes, thus eventually extending beyond the domain of validity of the equation.

As these problems with the two-dimensional theoretical results were identified, the need for a more accurate description than that provided by the nonlinear Schrödinger equation becomes apparent. A nonlinear integro-differential equation, first derived by Zakharov (1968), was identified by Crawford et al (1978) as being the desirable replacement. Zakharov's equation describes the time evolution of the complex envelope spectral function $\bar{A}(\mathbf{k},t)$,

$$i\frac{\partial}{\partial t}\bar{A}(\mathbf{k},t) = \iiint_{-\infty}^{\infty} T(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)\,\delta(\mathbf{k}+\mathbf{k}_1-\mathbf{k}_2-\mathbf{k}_3)$$

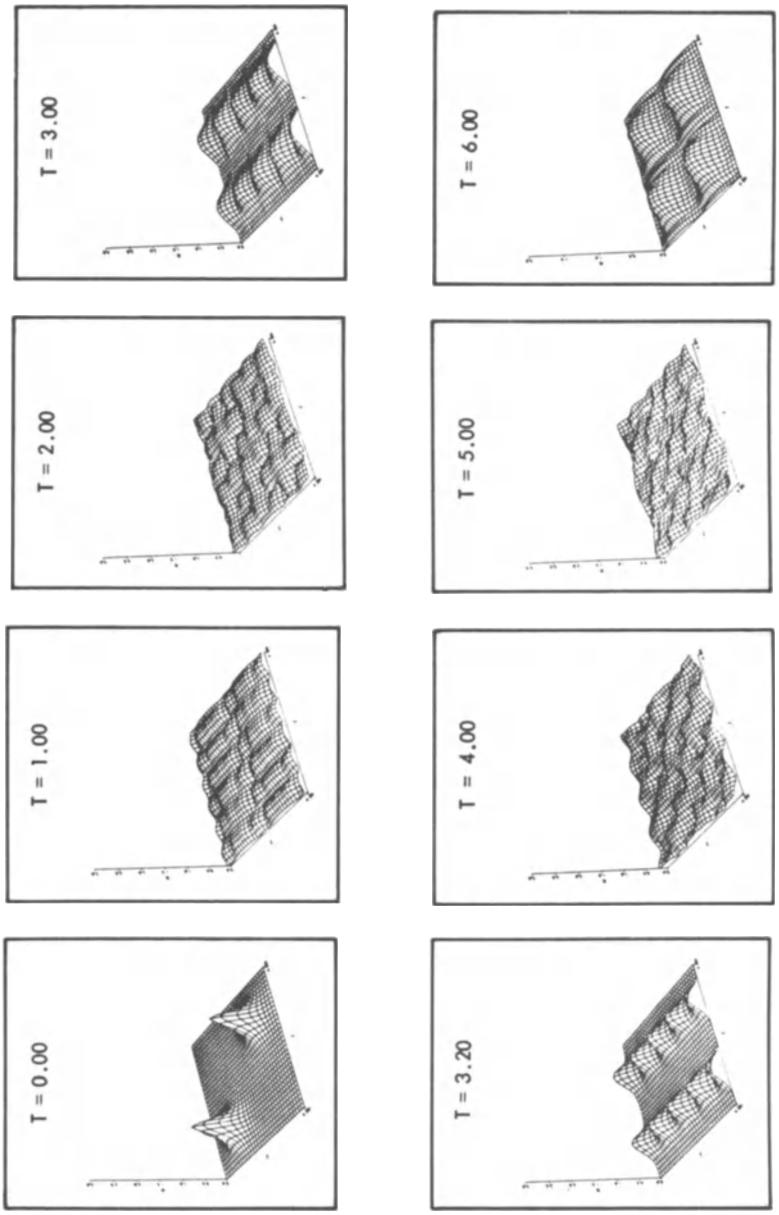
$$\times\, e^{i[\omega(\mathbf{k})+\omega(\mathbf{k}_1)-\omega(\mathbf{k}_2)-\omega(\mathbf{k}_3)]t}\,\bar{A}^*(\mathbf{k}_1)\bar{A}(\mathbf{k}_2)\bar{A}(\mathbf{k}_3)\,d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3 \qquad (18)$$

where $\bar{A}(\mathbf{k},t)$ is related to the free surface by the expression

$$\eta(\mathbf{x},t) = \frac{1}{2\pi}\int \frac{|\mathbf{k}|^{\frac{1}{4}}}{\sqrt{2}\,g^{\frac{1}{2}}}\left\{\bar{A}(\mathbf{k},t)e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} + \bar{A}^*(\mathbf{k},t)e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\right\}d\mathbf{k} \qquad (19)$$

and $T(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ is a real interaction coefficient first calculated by Zakharov (1968) with minor algebraic corrections made by Crawford, Saffman & Yuen (1980).

The nonlinear Schrödinger equation can be recovered, in one or two dimensions, by expanding the frequencies $\omega(\mathbf{k}_i)$ about a carrier-wave vector to second order, and replacing the interaction coefficient $T(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$ by its value when all four arguments are


T = 0.00
T = 1.00
T = 2.00
T = 3.00
T = 3.20
T = 4.00
T = 5.00
T = 6.00

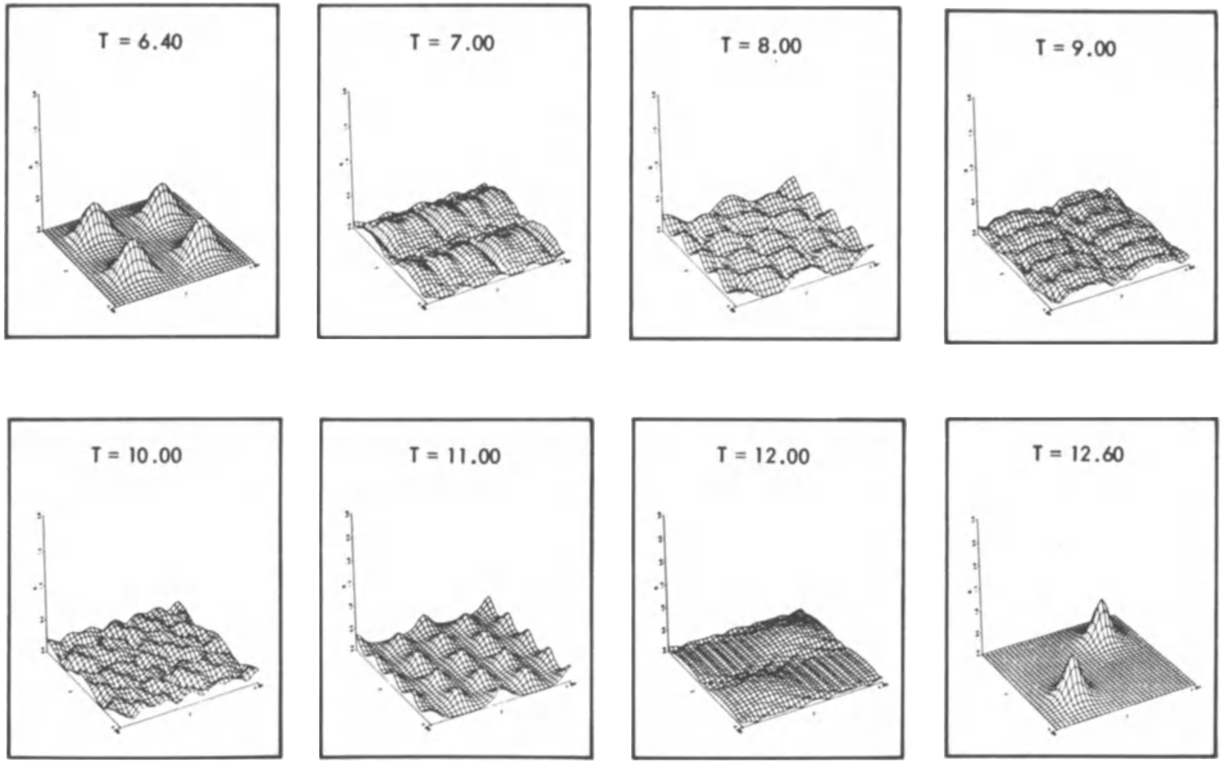

T = 6.40
T = 7.00
T = 8.00
T = 9.00
T = 10.00
T = 11.00
T = 12.00
T = 12.60

Figure 8. Example of recurrence in two space dimensions. Doubly periodic calculations of the two space dimensional nonlinear Schrödinger equation. Initial conditions with two concentrated packets evolve through complicated patterns but emerging unchanged in form at opposite locations.

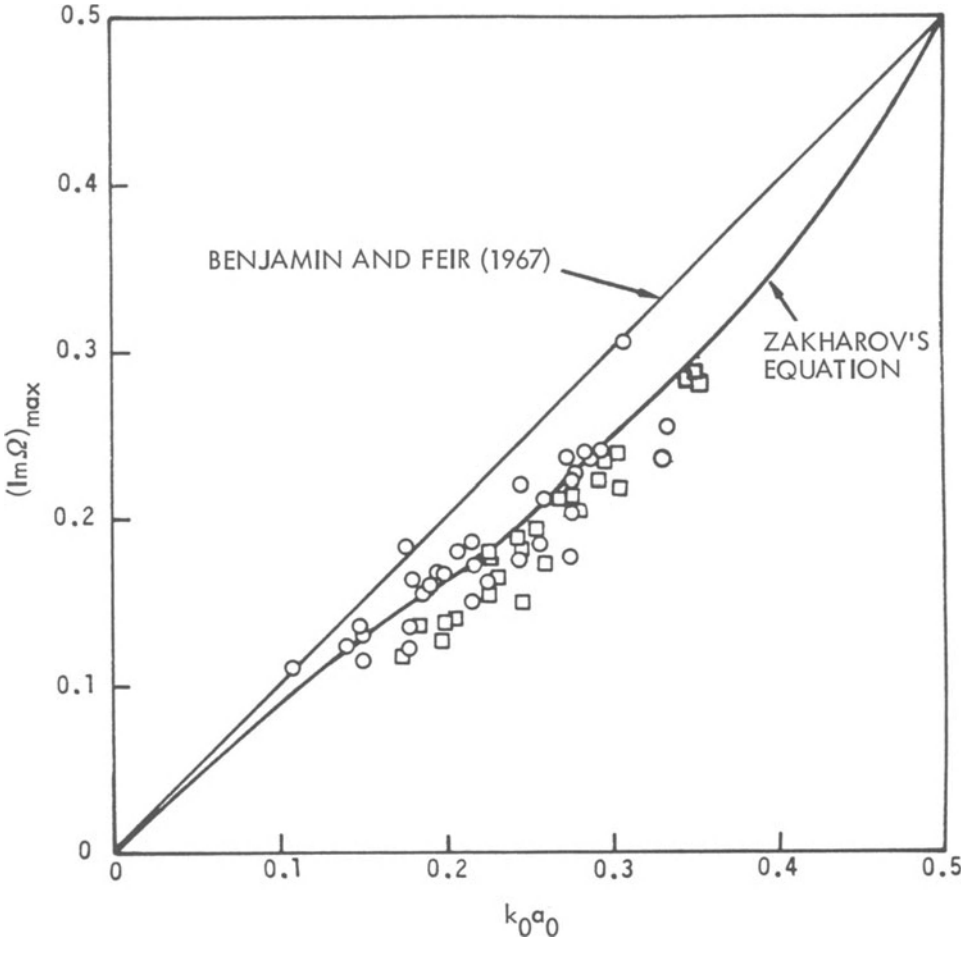


Figure 9. Plot of the most unstable modulational frequency $(\mathrm{Im}\Omega)_{\max}$ against carrier wave steepness k_0a_0. Experimental data taken from the TRW Fluid Mechanics Department experiment.

evaluated at the carrier-wave vector $\underset{\sim}{k}_o = (k_o,0)$. The resultant integral equation is then the Fourier-transformed equation for the complex amplitude A(x,y,t) that satisfies the nonlinear Schrödinger equation. This simplification is equivalent to retaining only the leading-order terms in nonlinearity and dispersion.

Results obtained from the Zakharov integral equation were found to be very encouraging. The stability analysis for one-dimensional modulational perturbation yields an instability growth rate lower than that predicted by the nonlinear Schrödinger equation (and by the Benjamin & Feir analysis). The difference becomes appreciable for moderate wave steepness (ka = 0.15 and higher), and the growth rate predicted by the Zakharov equation is in better agreement with experimental data (Crawford et al 1978; see Figure 9). For very large values of wave steepness, Zakharov's equation predicts a restabilization, which is in qualitative agreement with the exact numerical results of Longuet-Higgins (1978) and Peregrine & Thomas (1979).

The Zakharov equation can also be used to study the interaction of more than one wave component traveling in different directions. In particular, the problem of the generation of a third wavetrain by two resonantly interacting and intersecting wavetrains was examined by D. R. Crawford and H. C. Yuen (in preparation), and the results compare very well with the experimental data of McGoldrick et al (1966) and Longuet-Higgins & Smith (1966).

The most important result, however, is the stability analysis in two space dimensions. Crawford, Lake, Saffman & Yuen (1982) showed that, contrary to the results from the two-dimensional nonlinear Schrödinger equation, the instability diagram for the uniform wavetrain possesses a high-wavenumber cutoff (see Figure 10). Therefore the leakage to high modes experienced by certain solutions of the two-dimensional nonlinear Schrödinger equation need not occur in wave systems described by the Zakharov equation. The concept of wave coherence appears to be relevant to wave systems in two space dimensions after all, although positive confirmation has yet to be produced.

RESTABILIZATION OF THE UNIFORM WAVETRAIN AT LARGE AMPLITUDE

The modulation instability for weakly nonlinear wavetrains as calculated by Benjamin & Feir (1967), and from the nonlinear Schrödinger equation, is "similar" in the wave steepness (here taken to be $k_o a_o = \pi h/\lambda$), in the sense that when plotted in scaled variables, with the modulational wavenumber scaled by $k_o a_o$ and the growth rate scaled by $k_o^2 a_o^2$, the results are no longer dependent on

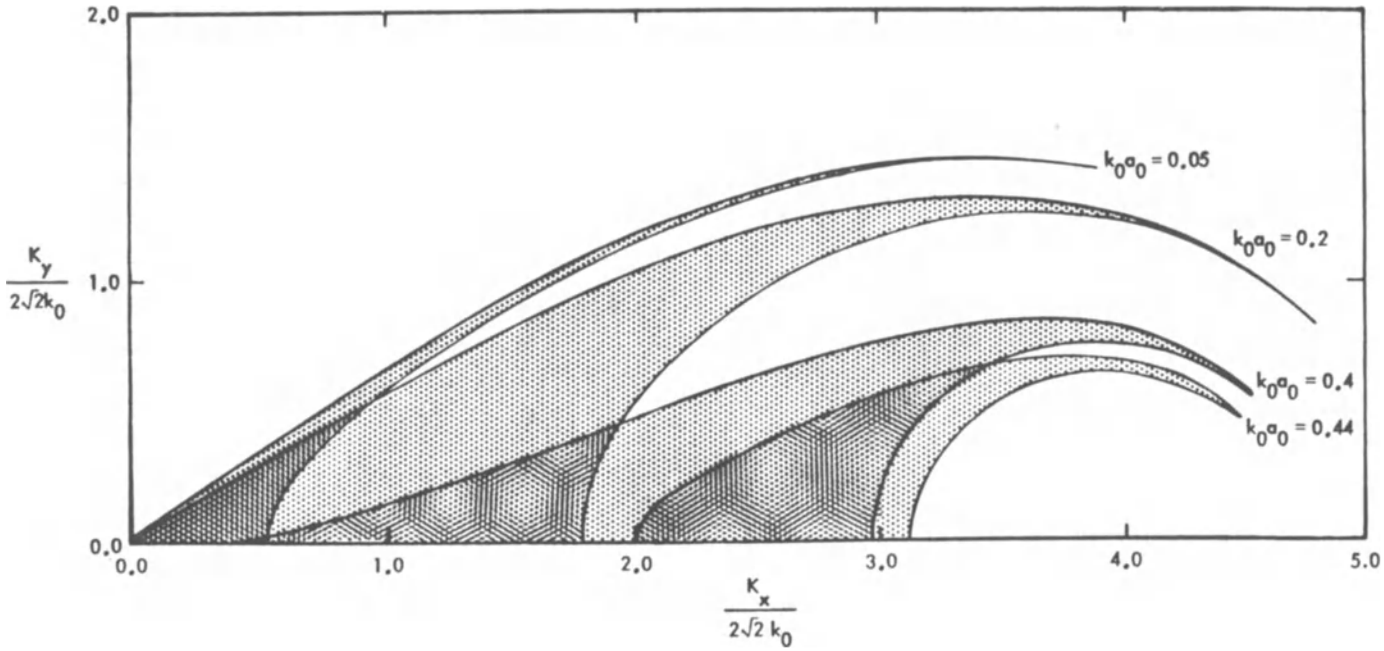


Figure 10. Stability boundary for a uniform wavetrain subject to infinitesimal two-dimensional perturbations based on the Zakharov equation. The shaded region represents instability. For k_oa_o greater than 0.495, the instability disappears and the wavetrain is restabilized.

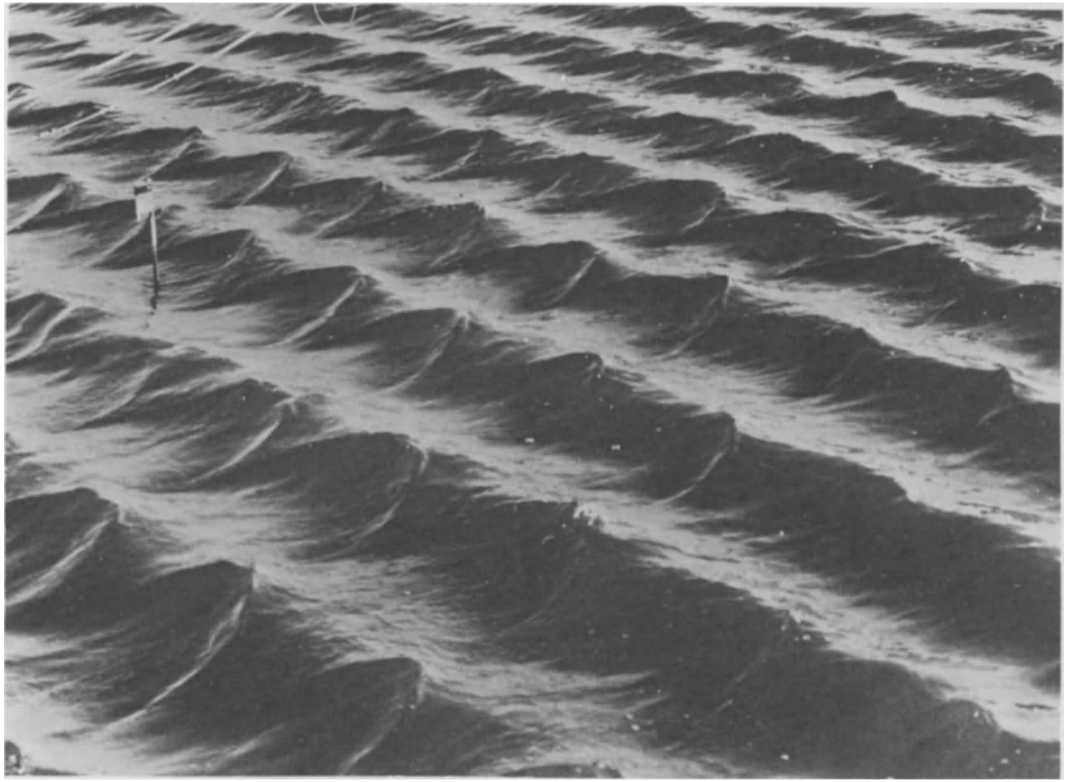


Figure 11. Experimental observation of steady three-dimensional wave patterns. This pattern results from a uniform wavetrain through the McLean et al. three-dimensional instability and compares well with the predictions of Saffman and Yuen. (M. Y. Su's experiment, NORDA)

k_0a_0. This is not true for the Zakharov equation, which has a stronger dependence on k_0a_0. In fact, the Zakharov equation indicates that the modulational instability disappears for sufficiently large values of k_0a_0. The Zakharov equation predicts wavetrain restabilization at $k_0a_0 \doteq 0.5$. Since this is greater than the k_0a_0 of the limiting wave, the prediction is only suggestive of the correct behavior. The value of k_0a_0 at which restabilization actually occurs is around 0.34, as shown by the exact numerical calculations of Longuet-Higgins (1978).

BIFURCATION

The restabilization of a wavetrain occurs when the eigenvalue of the perturbation equation changes from imaginary (representing growth and decay) to real (representing oscillations). Further increase in the wave amplitude results in the real eigenvalues becoming zero. When this occurs, it means that a disturbance can be put on the wave which neither grows nor oscillates. In a sense, it can become the starting point of a new branch of solutions and this point of zero eigenvalue can be a bifurcation point.

In the two-dimensional case, bifurcated Stokes waves were first discovered by Chen & Saffman (1980). They showed that steady modulated wave patterns with N waves per modulation period can exist at large wave amplitude. It is interesting to note that an earlier proof by Garabedian (1965) showed that a symmetric, uniform, deep-water, gravity wavetrain is unique provided that its crests and troughs are each of the same level. The findings of Chen & Saffman (1980) do not violate this theorem, but they do point to the crucial requirement of equal levels of crests and troughs in the uniqueness proof.

In the three-dimensional case, it can be shown (Saffman & Yuen 1980) that the critical wave steepness for bifurcation of oblique patterns decreases with increasing obliqueness. In fact, bifurcated solutions, in the form of symmetric and antisymmetric patterns, occur for every value of wave steepness. Calculations of these patterns, using the Zakharov equation and the exact water wave equations, were made by Saffman & Yuen (1980a,b) and Meiron, Saffman & Yuen (1982).

Independent of the theoretical predictions, Su (1982) reported experimental observations of regular, steady, three-dimensional wave patterns. The observed patterns compared well with the theoretical predictions of Saffman & Yuen, and provided one further piece of information: there is a preference for wave patterns with alternate high and low crests in the direction of propagation. The preference can be better understood with a subsequent calculation by McLean et al (1981) which uncovered a new type of three-dimensional instability which appears to trigger these bifurcation patterns (Figure 11).

NEW TYPE OF THREE-DIMENSIONAL INSTABILITY

McLean et al (1981) performed a stability analysis of a nonlinear wavetrain to infinitesimal three-dimensional perturbations using the exact water wave equations. They recovered the modulational instability, which for small wave steepness, corresponds to the Benjamin & Feir (1967) instability in the unidirectional case. They found, however, that for larger wave steepness, a higher order instability becomes prominent and eventually dominates. This new instability is strongly three-dimensional, with the maximum growth rate associated with disturbances corresponding to alternate wave crests being higher. In addition, perturbations corresponding to alternate wave crests being higher has the important property that they copropagate with the underlying wave. In other words, the instability will "lock on" to the wave and grow. This type of instability can cause the unstable uniform wavetrain to bifurcate into a wave pattern, and may explain why the most readily observed pattern is the one with alternate waves being higher. This new type of three-dimensional instability does not restabilize at large wave steepness, in contrast with the modulational instability mentioned earlier.

REFERENCES

Benjamin, T. B., Feir, J. E. 1967. The disintegration of wavetrains on deep water. Part 1. Theory. J. Fluid Mech. 27:417-30.

Chen, B., Saffman, P. G. 1980. Numerical evidence for the existence of new types of gravity waves of permanent form on deep water. Stud. Appl. Math. 62:1-21.

Chu, V. H., Mei, C. C. 1971. The evolution of Stokes waves in deep water. J. Fluid Mech. 47:337-51.

Crawford, D. R., Lake, B. M., Saffman, P. G., Yuen, H. C. 1981. Stability of weakly nonlinear deep-water waves in two and three dimensions. J. Fluid Mech. 105:177-91.

Garabedian, P. R. 1965. Surface waves of finite depth. J. Analyse Math. 14:161-69.

Gardner, C. S., Greene, J. M., Kruskal, M. D., Miura, R. M. 1967. Method for solving the Korteweg-de Vries equation. Phys. Rev. Lett. 19:1095-97.

Hasimoto, H., Ono, H. 1971. Nonlinear modulation of gravity waves. J. Phys. Soc. Jpn. 33:805-11.

Lake, B. M., Yuen, H. C., Rungaldier, H., Ferguson, W. E. Jr. 1977. Nonlinear deep-water waves: Theory and experiment. Part 2: Evolution of a continuous wave train. J. Fluid Mech. 83:49-74.

Lighthill, M. J. 1965. Contributions to the theory of waves in nonlinear dispersive systems. J. Inst. Math. Appl. 1:269-306.

Longuet-Higgins, M. S. 1978. The instabilities of gravity waves of finite amplitude in deep water. II. Subharmonics. Proc. R. Soc. London Ser. A 360:489-505.

Longuet-Higgins, M. S., Smith, N. D. 1966. An experiment on third order resonant wave interactions. J. Fluid Mech. 25:417-36.

Martin, D. U., Yuen, H. C. 1980. Quasi-recurring energy leakage in the two-space-dimensional nonlinear Schrödinger equation. Phys. Fluids 23:881-83.

McGoldrick, L. F., Phillips, O. M., Huang, N. E., Hodgson, T. H. 1966. Measurements of third-order resonant wave interactions. J. Fluid Mech. 25:437-56.

Meiron, D. I., Saffman, P. G., Yuen, H. C. 1982. Calculation of steady three-dimensional deep-water waves. J. Fluid Mech. 124:109-121.

McLean, J. W., Ma, Y. C., Martin, D. U., Saffman, P. G., Yuen, H. C. 1981. A new type of three-dimensional instability of finite amplitude gravity waves. Phys. Rev. Lett. 46:817-820.

Peregrine, D. H., Thomas, G. P. 1979. Finite-amplitude deep-water waves on currents. Proc. R. Soc. London Ser. A 292:371.

Saffman, P. G., Yuen, H. C. 1978. Stability of a plane soliton to infinitesimal two-dimensional perturbations. Phys. Fluids 21:1450-51.

Saffman, P. G., Yuen, H. C. 1980a. Bifurcation and symmetry breaking in nonlinear dispersive waves. Phys. Rev. Lett. 44:1097-1100.

Saffman, P. G., Yuen, H. C., 1980b. A new type of three-dimensional deep-water wave of permanent form. J. Fluid Mech. 101:797-808.

Saffman, P. G., Yuen, H. C. 1982. Three-dimensional deep-water waves. II. Calculation of steady symmetric wave patterns. Unpublished manuscript.

Stokes, G. G. 1849. On the theory of oscillatory waves. Trans. Cambridge Philos. Soc. 8:441-55. Math. Phys. Pap. 1:197-229.

Su, M. Y. 1982. Three-dimensional deep-water waves. Part 1. Experimental measurement of skew and symmetric wave patterns. J. Fluid Mech. 124:73-108.

Whitham, G. B. 1965. A general approach to linear and nonlinear dispersive waves using a Lagrangian. J. Fluid Mech. 22:273-83.

Whitham, G. B. 1967. Nonlinear dispersion of water waves. J. Fluid Mech. 27:399-412.

Whitham, G. B. 1970. Two timing, variational principles and waves. J. Fluid Mech. 44:373-95.

Yuen, H. C., Ferguson, W. E. Jr. 1978. Relationship between Benjamin-Feir instability and recurrence in the nonlinear Schrödinger equation. Phys. Fluids 21:1275-78.

Yuen, H. C., Lake, B. M. 1975. Nonlinear deep water waves: Theory and experiment. Phys. Fluids 18:956-60.

Zakharov, V. E. 1968. Stability of periodic waves of finite amplitude on the surface of a deep fluid. Zh. Prikl. Mekh. Tekh. Fiz. 1968, No. 2:86-94. Translated in J. Appl. Mech. Tech. Phys. 1968, No. 2:190-94.

Zakharov, V. E., Rubenchik, A. M. 1973. Instability of waveguides and solitons in nonlinear media. Zh. Eksp. Teor. Fiz. 65:997-1101. Translated in Sov. Phys. JETP 38:494-500 (1974).

Zakharov, V. E., Shabat, A. B. 1971. Exact theory of two-dimensional self-focusing and one-dimensional self-modulating waves in nonlinear media. Zh. Eksp. Teor. Fiz. 61:118-134. Translated in Sov. Phys. JETP 34:62-69 (1972).

BOSON CONDENSATION IN BIOLOGICAL SYSTEMS

G. Vitiello[1] and E. Del Giudice[2]
S. Doglia and M. Milani

[1]Istituto di Fisica dell'Universita -
84100 Salerno, Italy
INFN - Sezione di Napoli

[2]Dipartimento di Fisica dell'Universita, Via Celoria 16 -
20133 Milano, Italy

The accumulation of theoretical and experimental data about the components of a biological system is a fundamental step for understanding it. However a basic dynamical scheme is still lacking, actually a living system cannot be understood merely as the sum of the properties and behaviors of its components. This situation is very familiar in many body physics: a crystal, for example, is composed of atoms embedded in a lattice structure; but a complete description of the crystal can be obtained only if one considers other kinds of "particles", e.g. phonons, plasmons, etc, besides the original atoms. These new particles, which are the quanta of the system's collective excitations, appear rather peculiar. If one breaks up the crystal into its components, one does not find any trace of them -- one will obtain only the "original" atoms. The crystal is basically the "collective behavior" of its component atoms. Studying the crystal means studying the dynamical collective modes which are in fact those "particles" as the phonons, the plasmons, etc. In the same way a physicist's conclusion would be that a living system is characterized just by its dynamical collective modes. This is a formidable task because the basic dynamics might be very complicated. Such a task could be even out of the reach of any analytical computational scheme. Moreover a living system exhibits seemingly contradictory features.

It has been suggested that a living system has to be a dissipative one, namely it must be an open system able to dissipate outward all the incoming energy (1). From a macroscopic point of view this general principle of dissipativity must be compared with the equally general fact that the entropy of the system sharply decreases when it is metabolically active.

The decrease of entropy, i.e. some sort of ordering occurring in the system, would require an expense of energy by the thermodynamics laws. In our case we are then compelled to look for a dynamical mechanism able to order the system by a sheer transit of energy. One more puzzle emerges upon examination of the system biochemistry. The observed rigorous time-ordering of the chemical reactions occurring in the metabolic activity, without unwanted macroscopic fluctuations, occurs in the frame of a non-equilibrium thermodynamic behavior as required by the general principle of dissipativity.

Let us consider these difficulties with a little more detail. A possible approach could be to regard the macroscopic features of the biological system as a consequence of dynamical microscopic laws. This approach "works" in other fields of natural phenomena, so that we will assume that also in biology a macroscopic observable behavior would be the effect of a collective dynamics of elementary constituents submitted to a microscopic dynamics. In crystals the collective motions of atoms are described by means of microscopic "quasiparticles" such as phonons. Our aim is to find out the corresponding "quasiparticles" responsible for the behavior of a biological system.

The dynamics of a biological system will be complex enough to account for all the different observed behaviors. In order to get a first understanding of the most general features, we should concentrate on the symmetry properties of the system. This approach, which gets as much information as possible from the symmetry properties, without going through explicit dynamical computations, has been very successful in physics.

Quantum Field Theory (QFT) has proven to be suitable both in the description of the collective phenomena and for explaining the creation of ordered (not necessarily spatially-ordered) structures. For this reason we start from the QFT framework as the theoretical tool for understanding the dynamics of biological systems (2). Because living matter is basically a dielectric (3); we will try to describe it as a set of macromolecular dipoles, submitted to a Lagrangian of appropriate complexity. A dipole dynamics does not contain at a fundamental level any preferred direction, so that the Lagrangian, whatever its specific form, must be rotationally invariant, i.e. invariant under the SU(2) group. the macromolecule environment in a living system is water, whose polar molecule allows in principle aqueous electrets under suitable conditions. We take the aqueous environment of macromolecules as the "vacuum" of our system. Our QFT model then includes a rotationally invariant Lagrangian and a "vacuum", which could be characterized by a net electric polarization, implying a given space direction which breaks the rotational invariance.

Recently it has been recognized (4) that electrets (*not ferroelectrics*) consisting essentially of water molecules exist around the main biomolecules. These electrets couple to activated sites of the macromolecule and have an activation energy having an order of magnitude of one half of an electron volt. Moreover electrets have been shown to oscillate with frequencies as low as 10^{-3} to 10^{-5} Hz.

We can translate these experimental findings into a property of the vacuum in our scheme. The vacuum expectation value (v.e.v.) of the electric polarization field P may be nonvanishing:

$$\langle 0 \, | P(x) | \, 0 \rangle = P \neq 0 \tag{1}$$

where the direction of $\underline{P}$ will be taken as the 3rd direction.

The coexistence of an invariant Lagrangian and a non invariant vacuum is well-known in QFT. The system is governed by the Goldstone theorem (5). We will discuss briefly the topics in the framework of a specific formulation of QFT (6), useful for considering symmetry problems. This approach can be summarized as follows. We distinguish two levels of physical reality: the microscopic one and the phenomenological one. At the microscopic level the system is described by Heisenberg *interacting* fields (in our case the dipole fields of molecules), which obey nonlinear field equations:

$$\Lambda(\partial) \quad \chi(x,t) = J \, [\chi(x,t)] \tag{2}$$

where $\Lambda(\partial)$ is a kinetic operator and $\underline{J}$ a source operator. Eq. 2 is derived from a Lagrangian exhibiting a given degree of symmetry. At the phenomenological level, the system appears as a set of particles (its asymptomatic components in a LSZ formulation) or mutually non interacting quasi-particles (quanta of collective excitations such as phonons, plasmons, magnons and so on). They will be consequently described by operators ϕ, obeying free-field equations:

$$K(\partial) \; \phi = 0 \tag{3}$$

The Hilbert space of the physical states of the system is constructed from the solutions of Eq. 3. The matrix elements of the microscopic fields χ in this space cannot be expected to be particularly simple, so that any possibility of physical understanding is dependent upon the existence of a "dynamical map", namely a functional relationship between the matrix elements of the fields χ and the matrix elements of the fields ϕ:

$$< a|\chi(x,t)|b > = < a|\psi\ [\phi(x,t)]|b > \tag{4}$$

If the Lagrangian has a certain symmetry, i.e. the solutions of Eq. 2 are transformed one into another by a certain group of transformations, Eq. 4 will define a corresponding group for the fields ϕ, solutions of Eq. 3. If g is a transformation belonging to the invariant group of Eq. 4 so that $\chi' = g\chi$ and $\Lambda(\partial)\chi' = J[\chi']$ then Eq. 4 will define a $\phi' = h\phi$ such that

$$K(\partial)\ \phi' = 0$$

$$< a|\chi'|b > = < a|\Psi(\phi')|b > \tag{5}$$

The group of transformations h, defined by Eq. 5 and acting on the fields ϕ is not necessarily the same group of transformations acting on the fields χ. Suppose that some field ϕ transforms as $\phi(x) \to \phi(x) + cf(x)$ with $K(\partial)\ f(x) = 0$.

Then the transformation $\phi \to \phi + cf$ transforms one solution of Eq. 3 into another and belongs to the invariance group of Eq. 3. The dynamical map (4), maps this transformation into an element of the invariance group of Eq. 2. But the transformation $\phi \to \phi + cf$ just transforms an invariant vacuum into a non-invariant one, so it cannot be a member of the invariance group of Eq. 2 which only contains homogeneous transformations. The two invariance groups -- for Eqs. 2 and 3 respectively -- are proven to be different whenever the vacuum expectation value of some χ field is non-vanishing. They are mapped one into another by Eq. 4, which defines a dynamical rearrangement of the symmetry. The Lagrangian original symmetry gives rise to a different symmetry at the level of observable variables.

The occurrence of this rearrangement, as a consequence of the vacuum non invariance, implies that there exist fields which transform in a nonhomogeneous way. Moreover, due to the invariance of the theory -- this is the content of the Goldstone theorem (5) -- these fields are massless in absence of gauge fields.

Let us now come back to our original dipole system, characterized by an SU(2) invariant Lagrangian. The appearance of a water net polarization shown by Eq. 1, implies a rearrangement of the SU(2) invariance. It has been shown elsewhere (7) that the observable system is described by the group E(2), which is the "group contraction" of the original SU(2) (8,9). E(2) is a group formed by a rotation around an axis (in our case the polarization direction and two "translations" along the two normal directions.

The "translations" have the form:

$$\begin{aligned} B(x) &\longrightarrow B(x) + \text{constant} \\ B^{+}(x) &\longrightarrow B^{+}(x) + \text{constant} \end{aligned} \tag{6}$$

where $B(x)$ and $B^{+}(x)$ are the massless fields forseen by the Goldstone theorem. B and B^{+} preserve the dipole quantum number of the basic Heisenberg field χ, which would otherwise not be conserved because of the symmetry breaking. The massless fields B and B^{+} describe then electric polarization waves, that, because of the invariance transformations (6) must be also coherent and correspond to a Bose condensation of B-quanta.

In conclusion, the QFT machinery and the Goldstone theorem ensure that, given our hypothesis, a biological system must be characterized by coherent electric polarization waves. Frohlich (3) has predicted the existence of these waves under the condition of dissipativity, which then appears as the macroscopic manifestation of a microscopic invariance of the dynamics. We thus see how dissipativity is connected with the (spontaneous) entropy decrease; the energy release occurs through the production of a coherent wave.

Coherent polarization waves have been shown to be the consequence of a net water polarization. Where does this polarization comes from? It has been remarked recently (10) that long macromolecular chains may host slowly moving nonlinear excitations -- the solitons -- which are produced by the output of metabolic reactions. Such solitons can also capture electrons, giving rise to electric currents (11). It is not too unreasonable to assume that the fields associated with such currents could polarize the water dipoles. There is a numerical coincidence that looks interesting: The average energy to create a soliton is of the same order of magnitude of the activation energy for the electret. The electret then would appear as a consequence of the dynamical rearrangements occurring in the system, after the creation of solitons, without further expense of energy.

It has been shown (7) that Davydov solitons can be treated as localized Bose condensations of excitation quanta on the molecular one dimensional chains. This is actually a general result in QFT models (12, 13). It is then possible to outline a rough scheme (7, 14) for the dynamics of the biological process. The system first gets energy from the environment, not information, but just energy in an incoherent form. The organization of such energy on the molecular chains gives rise to local condensations. These spread out without expense of energy, producing water electrets. At this point the dynamics originally confined on the one-dimensional molecular chains invades all the system space, correlating previously separated domains through long range coherent electric waves. These waves play an important role in the architectural and biochemical organization of the system through the associated ponderomotive forces (15), as discussed in a different contribution to this conference.

Spectroscopic evidence (16) has been analyzed (14) in the framework of the above model.

In conclusion the QFT approach allows discussion of the collective behavior of a biological system in terms of spontaneous loss of symmetry and creation of order through Bose condensation of quanta. Much work has yet to be done in order to introduce a realistic finite size for the system, to consider temperature effects and so on. In ref. 7 preliminary comments on the finite size problem can be found together with a discussion on the system stability under soft external perturbations.

REFERENCES

1. I. Prigogine and G. Nicolis, "Self-Organization in Nonequilibrium Systems, from Dissipative Structures to Order through Fluctuations." Wiley, New York, (1977).
2. E. Schroddinger, "What is Life?" Cambridge, (1944).
3. H. Frohlich, Rivista del Nuovo Cimento, 7:399 (1977).
4. J.B. Hasted, H.M. Millany and D. Rosen, J. Chem. Soc.; Faraday Trans. 77:2289 (1981).; S. Celasxhi and S. Mascarenhas, Biophysical Journal 20:273 (1977).
5. J. Goldstone, Nuovo Cimento 19:154 (1961); J. Goldstone, A. Salam and S. Weinberg, Phys. Rev. 127:965 (1962).
6. H. Matsumoto, M. Tachiki and H. Umezawa, "Thermo-field Dynamics and Condensed States", North-Holland, Amsterdam (1982).
7. E. Del Giudice, S. Doglia, M. Milani and G. Vitiello, "A Quantum Field theoretical Approach to the Collective Behavior of Biological Systems", Preprint, (March, 1983).
8. N.N. Shah, H. Umezawa and G. Vitiello, Phys. Rev. 10B:4724 (1974).
9. C. De Concini and G. Vitiello, Nucl. Phys. 116B:141 (1976); C. De Concini and G. Vitiello, Phys. Lett. 70B:355 (1977).
10. A.S. Davydov, Biology and Quantum Mechanics, Pergamon, Oxford (1982); A.S. Davydov, Physica Scripta 20:387 (1979).
11. A.S. Davydov, Kiev, Preprint (1979).
12. H. Matsumoto, P. Sodano and H. Umezawa, Phys. Rev. 19:511 (1979).
13. L. Mercaldo, I. Rabuffo and G. Vitiello, Nucl. Phys. 188B:193 (1981).
14. E. Del Giudice, S. Doglia and M. Milani, Physica Scripta 26:232 (1982).

15. E. Del Giudice, S. Doglia and M. Milani, *in*: "Coherent excitations in Biological systems", H. Frohlich and F. Kremer (eds), Springer, Berlin, p. 147 (1983).
16. S.J. Webb, *Phys. Rep.* *60*:201 (1980).

ORDER AND STRUCTURES IN LIVING SYSTEMS

E. Del Giudice[1], S. Doglia[2] and M. Milani[3]

[1]Dipartmento di Fisica, Universita di Milano, Via Celoria 16, I-20133 Milano, Italy; Istituto Nazionale di Fisica Nucleare, Sezione di Milano

[2]Gruppo Nazionale Struttura della Materia del C.N.R., Milano

[3]Istituto Nazionale di Fisica Nucleare, Sezione di Milano

INTRODUCTION

Biological systems are characterized by structure and dynamical ordering. Molecular biologists have gone a long way towards understanding structure. Research into dynamics is in progress. The aim is to elucidate how the large number of molecules in a cell can arrange themselves in ordered and meaningful patterns, avoiding mistakes and the kind of disordered behavior which produces heating. Actually the constance of temperature during the biological life cycle can be considered as a rough principle of biology. This means from a physical point of view that the energy released from metabolic reactions supplies collective degrees of freedom without any thermal losses, i.e. without increasing the system temperature.

Order in biological systems is not static as in crystals but appears in different forms at different times of the life cycle and reshuffles itself following environmental perturbations. A flexible order emerges as a feedback in the interaction between the system and the environment. This kind of order cannot be produced by a simple spatial ordering of components as in crystals, but needs long-range time-dependent phase correlations among the components. Living matter is a dielectric; consequently the above correlation must be connected to a coherent electric vibration of appropriate frequency. Herbert Fröhlich (1-3) was the first to stress this point; over the years he worked out a theory of such waves predicting a number of consequences. A dynamical model has been presented in these proceedings (4,5) which shows how coherent polarization waves could emerge as a consequence of the dynamics of biomolecules and water triggered by metabolism. In this model, early dynamics produce solitons on molecular chains. Subsequently surrounding water is driven to an electret state, thus preparing the necessary conditions for the onset of a collective regime dominated by coherent electric polarization waves.

The above philosophical point of view can be verified experimentally. Phase ordering in a material can be actually detected by Raman spectroscopy, (6) if the frequency of the phase correlation waves is in the appropriate range. The Raman scattering, i.e. the inelastic scattering of coherent electromagnetic radiation upon a collection of scattering centers preserves the phase relationships existing inside the collection. The emerging radiation then will have the same degree of coherence as the target oscillators. A completely disordered material provides a Raman response characterized by very large bands in a structureless background while sharp lines spectra are characteristic of phase ordered materials (7). This phase ordering might have a twofold origin: either could be a consequence of an extended spatial order as in crystals or could be produced by a coherent oscillation of given frequency existing in the medium.

Sydney J. Webb (6) has studied the Raman response of biological materials both in metabolic and in anabiotic conditions. A striking pattern emerges from his investigation. Sharp spectral lines are present at room temperature in metabolic conditions and are time dependent in a reproducible way. Metabolic inactivity corresponds to an empty spectrum. Under the above considerations, we expect a time dependent order in metabolically active biological systems which disappears when metabolic activity is stopped.

A purely spatial order as in solids or in liquid crystals cannot be the general answer. We cannot exclude that such order could hold in particular conditions, but the global consideration of biological phenomenology brings us to exclude the presence of a purely spatial order in living matter. Rather, we conclude the evidence supports the hypothesis of a dynamical order induced by coherent oscillations.

At this point we are confronted with a major difficulty. Living matter is made of up to 70% water and we know that water strongly absorbs radiation in a wide spectral range which includes the region of the Raman lines detected by Webb. Polarization waves have been introduced as a model for order at very long distances, but in usual conditions water would prevent them from traveling over an appreciable range. The question is whether a mechanism exists which could allow a polarization wave to travel a long distance inside "hostile" medium. From long range communication technology we recieve a positive answer; the means is by a waveguide.

In the biological case the waveguide has to be a part of the whole system, sharing with it the same requisites of flexibility and time dependence; then it must be produced by the same dynamical process responsible for the polarization waves. If the electric field $\vec{E}$ associated with the polarization wave is intense enough (say about 10^7 V/m) the wave propagation is governed by a nonlinear law which replaces the usual linear approximation responsible for the well known ordinary optics. Actually it has been found that electric fields in biological systems can reach very high values of the order of 10^7 to 10^8 V/m,(8) corresponding to a power in vacuum of about $5 \times 10^{11} \div 5 \times 10^{13}$ W/m^2.

We recall that the atomic electric fields are about 10^9 V/m in semiconductors and about 10^{11} V/m in dielectrics and therefore the existence of such high electric fields in biological systems implies that the nonlinear response has to be taken into account.

1. SELF-FOCUSING AND PONDEROMOTIVE FORCES OF COHERENT ELECTRIC WAVES

It is therefore important to investigate the role of nonlinearities in the propagation of intense coherent electric waves in biological media.*

We will analyze this propagation both in the intracellular case, where the medium is the cytoplasm (cytosol) and in the intercellular one, where in the particular case of erythrocytes (red blood cells) we consider the medium is the serum with blood proteins. The medium in which the electric wave

$$\vec{E} = \frac{1}{2}\vec{A}(r,t)\ \exp i(\omega_o t - \vec{k}_o \cdot \vec{r}) \quad + \text{c.c.} \qquad (1)$$

is propagating can be considered to be a homogeneous aqueous solution of macromolecules. It can be macroscopically described by its refractive index n, which is a function of the electric field E

$$n = n_o + n_2|A|^2 + n_4|A|^4 + \ldots \qquad (2)$$

---

*Editor's note: Whether such fields really exist in biological systems is problematic. The line widths and power levels necessary for nonlinear effects to become significant seem to be incompatible with the disorder of these systems. See also Bullough's contribution, these proceedings.

where n_o is the linear index of refraction (the asymtotic value of n as $\vec{A} \overset{o}{\to} 0$). This means that the refractive index of the medium is different in presence or in absence of an electric field and that in the field of a bounded light beam a medium which is initially homogeneous becomes optically inhomogeneous. In fact

$$n = n_o + n_2|A|^2 + n_4|A|^4 + \ldots.$$

in the region traversed by the beam and $n = n_o$ outside.

Let us confine ourselves to the lowest order term in nonlinear refraction; when $n_2 > 0$, the region where the field presents a maximum intensity is also the most optically dense one. A nonlinear refraction is usually produced in the Kerr effect, leading to the appearance of birefringence induced by the electric field. We will discuss this point in detail in the following.

We consider now the effect of a nonlinear index of refraction with $n_2 > 0$ on a beam propagating in the medium. It can be shown (9-12) that under proper conditions associated to the power of the electric field and to peculiar characteristics of the medium, self-focusing of the beam can take place, i.e. the propagating beam shrinks until a limiting diameter is reached after a path of critical length

$$R_{nl} = \frac{a}{2}\sqrt{\frac{n_o}{n_2|A|^2}} \tag{3}$$

a being the radius of the incoming beam. The quantity R_{nl}, which has the dimension of a length, is called the effective self-focusing length. When nonlinear refraction completly compensates for the diffraction spreading, the beam is confined within an optical waveguide ("filament") where it propagates without divergence and whose length depends on the field intensity. This process is usually called "self-trapping" of the wave. In this regime the beam is able to concentrate its energy on very small cross sections and the penetration range is greatly enhanced.

Self-focusing and self-trapping of the beam occurs only when a critical power P_{cr} is exceeded

$$P_{cr} = \frac{(1.22)^2\lambda^2 c}{256\ n_2} \tag{4}$$

as can be seen by simple arguments of geometrical optics. These properties can be found in any transparent medium exhibiting relevant Kerr nonlinearities. A large experimental evidence has been accumulated for many substances such as CS_2, nitrobenzene, CCl_4, etc. (11-13). It is interesting to note that theoretical estimates and experimental evidence of self-focusing has also been given for water, (10) with a critical power of the order of 1 MW to be compared with 0.2 MW for CS_2 and 0.25 MW for benzene. Let us now consider the dynamical consequences of a self-trapped beam propagating in an aqueous solution of biomolecules, i.e. in a medium of biological interest (14,15). The unit volume force (16) exerted by an electric field $\vec{E}$ in a medium with polarization. $\vec{P}$ is

$$\vec{F} = \rho\vec{E} + (\vec{P}.\vec{\nabla})\vec{E} + \frac{1}{c}\frac{\partial\vec{P}}{\partial t} \times \vec{H} \qquad (5)$$

where is the charge density, $\vec{E}$ and $\vec{H}$ the electric and magnetic fields of the wave, c the light speed. The linear term describes the interactions due to charge distributions, the second term (the "gradient" force) is very important on the boundary around the self-trapped beam and the third term is the radiation pressure term (17). It can be shown (18) that the gradient term in eq. 5 can be written as

$$(\vec{P}.\vec{\nabla})\vec{E} = \text{const.} \sum_k \left\{\frac{\omega_{ok}^2 - \omega^2}{(\omega_{ok}^2 - \omega^2)^2 + \Gamma_k^2}\right\} \nabla E^2 \qquad (6)$$

in terms of ω_{ok}, a resonance frequency of the dielectric and of Γ_k, the damping constant of the k-th oscillator. It is apparent that the gradient force plays an important role at the boundary of the self-trapped beam, because in this case the factor ∇E^2 is highly enhanced with respect to the gradient of the square of the incoming electric beam (up to a factor $10^2 - 10^3$).

It is also apparent from the structure of eq. 6 that this force will act selectively on different solute molecules according to the resonance properties between typical frequencies of the molecules and the frequency of the propagating beam. This force, directed along ∇E^2 and therefore normal to the axis of the beam, is attractive (or repulsive) depending on the sign of $(\omega_{ok} - \omega)$. More precisely molecules will be attractedd toward the beam when $\omega < \omega_{ok}$ and repulsed away from the beam when $\omega > \omega_{ok}$. As a consequence, condensation or rarefaction of molecules in the solution, i.e. concentration changes, around the "filament" can take place when ω approaches ω_{ok}. At this point it is important to recall that (13) the time scale of this process (electrostriction) is of the order of 10^{-8} s which has to be compared with Kerr mechanism characteristic time (10^{-11} to 10^{-12} s). Therefore, the propagating self-trapped beam might construct a material waveguide in the medium, provided that the coherent electric signal would last longer than the electrostriction characteristic time. The formation of the material filament requires the following conditions: i) high Kerr nonlinearity, to allow the presence of high electric field gradients on the boundary of the beam and its self-focusing; ii) electric field power greater than P_{cr} in order to stabilize self-focusing into self-trapping of the beam, and provide an efficient propagation of the signal in the medium; iii) resonance or quasi-resonance between the frequency of the electric wave and some frequencies of the molecules present in the medium in order to concentrate them around the beam; iv) amplitude A of the signal E constant over a time interval long on the electrostriction time scale (10^{-8} s) in order to build up the material waveguide around the filament.

Conditions i) and ii) will then ensure an efficient propagation of the wave overcoming the usual geometric and diffraction losses. Condition iii) and iv) on the frequency of the wave will allow (for $\omega < \omega_{ok}$) the formation of a material filament built up by selected molecules. A filament, in turn, will enhance the efficiency of the propagation and will screen the beam from external perturbations.

To establish whether efficient propagation of an electric wave can take place, Kerr properties of the medium have to be considered. Kerr constants at the frequency ω of the wave should be examined. The a.c. Kerr constants in the visible range for CS_2, nitrobenzene, and CCl_4 are of the order of 10^{-7} cm/ (Stat V)^2 (12). We are not aware of a.c. Kerr measurements for solutions of biomolecules. On the other hand, it can be seen that the knowledge of d.c. Kerr constants gives useful information on the a.c. Kerr response (10). The polarization of the medium induced by an a.c. electric field E is in a simple scalar form:

$$P = P(E) = \alpha E + \chi E^2 + \theta E^3 + \ldots\ldots$$

This formula contains both a static contribution and an alternating one at the field frequency (20). For a solution of polar molecules the d.c. contribution (which can be negative) to the Kerr constant dominates the a.c. one. Therefore the Kerr response of the medium will be approximated fairly well by its d.c. Kerr response. In the case of non-polar molecules, the static and alternating contributions are instead of the same order of magnitude and therefore again the Kerr response of the medium will be qualitatively given by its d.c. Kerr constant (10).

On the basis of these considerations, a knowledge of d.c. Kerr constants allows an estimation of the likelihood of self-focusing. For negative Kerr birefringence ($n_2 < o$) a defocusing of the beam occurs. For positive Kerr birefringence ($n_2 < o$) self-focusing and self-trapping of the beam can take place provided that the beam power exceeds P_{cr}.

In standard experiments (12) on self-focusing in CS_2, nitrobenzene, CCl_4 the beam power is of the order of 10 ÷ 100 KW, which is consistent with intensity of electric fields of 10^7 V/m. Since $P_{cr} \sim \frac{1}{n_2}$ and $n_2 = n_2(c)$ where c is the concentration of the solute molecules, a strong dependence of P_{cr} and therefore of self-focusing on the concentration of molecules are expected. A necessary condition for self-association and polymerization is then fulfilled and in presence of proper conditions (ordering among the monomers, (21) bridging agents, etc.) a polymer can be formed along the filament (15).

To complete the description of the full dynamics of the filament, it is necessary to discuss the last term in eq. 5, namely the radiation pressure term (22). A transversal electric wave produces a force directed along the propagation axis which is able to push molecules along the waveguide. A consequence is a pressure on the monomers attracted by the gradient force and ordered by the electric field in the filament. This pressure could actually play a role in providing the energy for a possible polymerization.

Moreover, the radiation pressure could account for the dynamics of the polymer, when assembly and disassembly at opposite ends of the filament take place, together with movement of monomers along the polymer itself. This process (treadmilling) is discussed in the next section when actin polymerization is examined.

If self-focusing of coherent electric waves and the consequent ponderomotive force arising in the biological medium are responsible for the polymer formation, then the critical length of the polymer could be related to the length of the self-trapped beam.

The interplay between gradient force and radiation pressure has been observed in physical systems by laser induced "levitation" and laser isotope separation (25). This "optical mechanics", while ensuring the propagation of the coherent waves, could actually provide the engine for transport mechanism and selected molecules interactions inside biological systems.

2. ORGANIZATION OF LIVING MATTER THROUGH COHERENT ELECTRIC WAVES

The above mechanisms of self-focusing and self-trapping together with the associated ponderomotive forces may support the efficient propagation of coherent electric waves through living matter.

We may compare the consequences of such a mechanism with the biological evidence.

Cytoskeleton Structures - Cell cytoplasm has been recently recognized as a complex structure, resulting from three networks of branched filaments of different diameters, whose interstices are filled by water (26). These three networks of filaments are made up of different proteins: i) microfilaments formed by actin whose diameter is about 60 Å, ii) microtubules made up from tubulin whose diameter is about 250 Å, and iii) intermediate filaments formed by a number of different proteins whose diameter is about 100 Å (27).

It has been shown that these microstructures are connected with metabolic activity and their shape and size depend on the cell life cycle and the interaction of the cell with the environment (26-29). It has also been suggested that the bulk of the cell chemical reactions occurs along these networks connecting all the principal cell organelles. The system of filaments has not been found to be directly attached to cell membranes. In presence of external electric and magnetic fields the whole filament network has been found to rotate, suggesting that filaments are not rigidly bound to the cell membrane. The actin network seems to start from selected sites in the proximity of the membrane.

All these filaments exhibit a constant and rapid turnover of their protein monomers (28-30). Moreover, it has been observed experimentally in vitro (29,30) that tubulin and actin filaments can acquire monomers at one end while loosing monomers at the opposite end, producing an associated movement of monomers along the filament. This process is called treadmilling and has been proposed to be present in vivo.

Our model of the formation and the dynamics of cytoskeleton is a consequence of the Fröhlich theory of coherent electric waves and of the nonlinear propagation of these waves within the cell cytoplasm (32). The existence of high electric fields of the order of 10^7 V/m on the cell membrane together with the nonlinear response of the cell medium, will ensure self-focusing and self-trapping of coherent electric waves produced by collective dynamics in the membrane (1-3).

In order to evaluate the nonlinear response of the cell medium, which can be treated as a solution of the cytoskeleton proteins, we may estimate Kerr properties. Positive birefringence (d.c. Kerr effect) has been observed at rather low values of external electric fields (E 10^4 V/m) for solution of G-actin and small polymers (of length $\lesssim 1.5\ \mu m$) of F-actin (33). Actin, composing about 20% of the total cell proteins, is therefore responsible for the nonlinear behavior of the cell cytoplasms. The efficient propagation of the coherent waves within the cell, as discussed in section 1, gives rise to ponderomotive forces which can build up a material filament network. As a consequence of changes in concentration in the medium along the filament, self-association of monomers and eventually polymer formation favored by ordering processes, can take place in strict dependence upon dynamical events in the whole cell. This would meet the requirements of J. Hyams for the cytoskeleton as "a vectorial framework upon which some other force generating component is organized" (34).

An interesting implication of this model is that a critical length R_{nl} (see eq. 3) has to be traveled in the medium before self-trapping of the wave occurs. Then an empty space is expected to show up between the starting filament and the source (e.g. cell membrane) of the coherent waves. Experimental evidence of this behavior would give further support to the proposed model.

Moreover the treadmilling process, reported above, could be accounted for by the interplay of the radiation pressure and of the gradient term in the ponderomotive force.

Blood Clotting and Contractils Formation - Erythrocytes, red blood cells of 8 μm diameter, are easily involved in adhesion processes; they form rouleaux that appear under the microscope as stacks of coins. Rouleaux formation, a typical process in blood clotting, is known to increase abnormally in a number of diseases such as coronary heart disease.

This process cannot be explained from the theory of colloid aggregation based on Brownian motion (35). It has been reported (36) that "abnormal" adhesion among erythrocytes (suspended in their own plasma) is observed unless: 1) membrane potential is suppressed; 2) metabolic activity is depleted; 3) erythrocyte membrane is disorganized by interfering agents altering the bilayer structure (glutaraldehyde). Furthermore, experiments on rouleaux formation among cells of different mammalian species show that this interaction is species specific. These results, again, cannot be accounted for from the theory of polydispersed colloid aggregation, but can be explained from Fröhlich's theory of coherent electric waves and long range interactions (36).

More recently (37) it has been found that rouleaux formation depends also upon the presence in the suspending medium of proper macromolecules and it does not occur in serum, i.e. in blood plasma depleted by proteins. Fibrinogen, PVP, dextran and polyethylene oxide, when added at proper concentrations to an isotonic saline suspension of erythrocytes, induce efficient rouleaux formation. No rouleaux formation occurs when albumin is added, even at concentrations greater than the physiological one. Electron microscopy studies have shown that contractile fibrils are present among adjacent erythrocytes when they are pulled apart. Their length ranges from few m to more than 100 μm, for erythrocytes of the same species. Much shorter and weaker contractils have been observed upon reduction of membrane potential and metabolic activity and in mixed interspecies rouleaux.

It has also been observed that the length of contractils reduces to zero when traction among cells forming rouleaux is released. This result cannot, therefore, be explained by the viscoelastic property of fibrils.

We believe that contractils are produced by Fröhlich activity of the erythrocytes and the propagation of coherent electric waves in the suspending medium. The question is whether self-focusing and self-trapping of the waves can take place and be responsible for contractils formation in the presence of proper macromolecules. For this purpose, Kerr properties of fibrinogen and albumin can be compared. In the case of fibrinogen solutions (at physiological pH) positive birefringence with high d.c. Kerr constant has been observed in presence of external electric field of the order of 10^6 V/cm (38). For albumin, a d.c. Kerr constant two orders of magnitude lower than that of fibrinogen is reported (39). Self-focusing and self-trapping of Fröhlich waves among erythrocytes could then be at work in presence of fibrinogen, with consequent formation of contractils as discussed in Section 1.

Also the existence of a threshold concentration of the rouleaux-inducing molecules in the medium can be explained from the model. The critical power for self-trapping of the coherent waves is a function of the refractive index of the medium and therefore of the concentration of the molecules.

A striking experimental characteristic of contractils is the presence of a small gap at each end of the fibrils, i.e. an empty space between erythrocyte and its formed contractil (37). The existence of such gaps is a critical requirement of the proposed model. As discussed above, self-trapping of the waves takes place after a path of critical length R_{nl} (see eq. 3). Along this critical path the gradient term of the ponderomotive force is then not strong enough to attract molecules and therefore no filament can be found.

It will be then very interesting to investigate experimentally the Kerr response of the other rouleaux inducing molecules, PVP, dextran and PEO.

In conclusion, the hypothesis of the existence of coherent electric vibrations in biological systems allows us to predict a number of phenomenological consequences and to explain critical experimental features.

REFERENCES

1. H. Fröhlich, Nature 228:1093 (1970).
2. H. Fröhlich, Adv. Electron. Electron. Phys. 53:85 (1980).
3. H. Fröhlich, Int. J. of Quantum Chemistry, vol. XXII (1983).
4. E. Del Giudice, S. Doglia and M. Milani, Physica Scripta 26:232 (1982).
5. E. Del Giudice, S. Doglia, M. Milani and G. Vitiello, a) to appear in Phys. Lett. A; b) submitted to Nuclear Physics B(FS).
6. S.J. Webb, Phys. Rep. 60:201 (1980).
7. E. Del Giudice, S. Doglia, M. Milani and M. Fontana, (in preparation).
8. H.T. Witt, Biochim. Biophys. Acta 505:355 (1979).
9. G.A. Askar'yan, Sov. Phys. Usp. 16:680 (1974).
10. R.Y. Chiao, E. Garmire and G.H. Townes, Phys. Rev. Lett. 13:479 (1964).
11. Y.R. Shen, Phys. Lett. 20:378 (1966).
12. S.A. Akhmanov, A.P. Sukhorukov and R. Khokholov, Sov. Phys. Usp. 93:609 (1967).
13. O. Svelto, Progress in Optics E. Wolf (ed.) XII, North Holland, p. 1(1974).
14. E. Del Giudice, S. Doglia and M. Milani, Phys. Lett. 85A:402 (1981).
15. E. Del Giudice, S. Doglia and M. Milani in: "Coherent Excitations in Biological Systems" H. Fröhlich and F. Kremer (eds.) Springer, Berlin, p. 123 (1983).
16. A. Einstein and J. Laub, Ann. Phys. (leipz.) 26:541 (1908).
17. A. Ashkin, Phys. Rev. Lett. 24:156 (1970).
18. G.A. Askar'yan, Sov. Phys. JETP 15:1088 (1962).
19. G. Mayer and F. Gires, Compt. Rend. 258:2039 (1962).
20. I. Tarassov, "Bases Physiques de l'Electronique Quantique" (Editions Mir, Moscow) (1979).
21. T. Folda, L. Gros and H. Ringsdorf, Makromol. Chem., Rapid Comm. 3:167 (1982).
22. A. Ashkin and J.M. Dziedzic, Appl. Phys. Lett. 19:283 (1971).
23. N.V. Karlov and A.M. Prokhorov, Sov. Phys. Usp. 19:285; N.G. Basov, E.M. Belenov, V.A. Isakov, E.P. Markin, A.N. Oraevskii and V.I. Romanenko, Sov. Phys. Usp. 20:209 (1977).
24. F.K. Buckley and K.R. Porter, Protoplasma 64:349 (1967).
25. J.S. Klegg, Collect. Phenomena 3:162 (1981)a; J.S. Clegg, in: "Coherent Excitations in Biological Systems", H. Fröhlich and F. Kremer (eds.), Springer, Berlin, p. 189 (1983)b.
26. U. Lindberg, L. Carlsson, F. Markey and L.E. Nystrom, Meth. Achiev. Exp. Pathol. 8:143 (1979).
27. A. Weeds, Nature 296:811 (1982).
28. R.L. Margolis and L. Wilson, Nature 298:705 (1981).
29. M.P. Sheets and J.A. Spudich, Nature 303:31 (1983).

30. E. Del Giudice, S. Doglia and M. Milani, a) Phys. Lett. 90A:104 (1982); b) in: "The Application of Laser Light Scattering to the Study of Biological Motion", J.C. Earnshow and M.W. Steer, (eds.) Plenum Press, New York, p. 493 (1982).
31. S. Kobayasi, H. Asai and F. Oosawa, Biochim. Biophys. Acta 88:528 (1964).
32. J. Hyams, Nature 291:107 (1981).
33. M.U. Smoluchowski, Zeit. Phys. Chim. 92:129 (1917).
34. S. Rowlands, L.S. Sewchand, R.E. Lovlin, J.S. Beck and E.G. Enns, Phys. Lett. 82A:436 (1981).
35. S. Rowlands in: "Coherent Excitations in Biological Systems", H. Fröhlich and F. Kremer (eds.) Springer, Berlin, p. 145 (1983).
36. I. Tinoco, J. Am. Chem. Soc. 77:3476 (1955).
37. E. Fredericq and C. Houssier in: "Electric Dichroism and Electric Birefringence", Clarendon, Oxford (1975).

APPLICATIONS OF NONLINEAR PHYSICS IN BIOPHYSICS, BIOCHEMISTRY AND CELL BIOLOGY

GENERAL THEORY OF COHERENT EXCITATIONS ON BIOLOGICAL SYSTEMS

H. Fröhlich

Department of Physics
Oliver Lodge Laboratory
The University of Liverpool Oxford Street
Liverpool L69 3BX ENGLAND

ABSTRACT

Coherent excitations, as first proposed in 1968, can arise in the presence of nonlinear interactions if metabolic energy is supplied at a rate s above a critical s_o. Three types of excitations will be shown to exist: A excitation of a single mode; B excitation of a metastable highly polar state; C excitation of limit cycles or Lotka-Volterra oscillations in complex systems. They may have far reaching consequences on the dynamics of biological systems.

Externally applied radiation may, at certain frequencies and at certain stages of development, act as trigger to biological events. In this case a steplike dependence on the intensity is to be expected.

Relevant experimental evidence will be indicated.

INTRODUCTION

A general theory of biological activity necessarily will be concerned with properties common to most biological materials. From the point of view of physics the extraordinary dielectric properties (1) are outstanding in this respect. Most remarkable amongst these are the very high electric fields of about 10^7 volt/m, usually present in biological membranes, leading to very strong polarization of the materials in them.

Biologically the term "life-cycle" already indicates a crude periodicity in biological dynamics, though here, each particular case exhibits individuality prohibiting the exact reproducibility of a period.

We shall take a vast step and assume that coherent excitation of electric vibrations in a large range of frequencies play an important role in biological activity. Modern physics offers this type of excitation arising through nonlinear interactions, with the necessary energy supplied, in general, from metabolic processes. Thus while supply of energy to a material usually leads to heating, coherent excitations may arise if certain conditions are satisfied.

This point of view was first presented more than fifteen years ago (2,3). It has led to far reaching consequences which stimulated a great number of experiments (review up to 1979 in (4). In a recent meeting on coherent excitations in biological systems (5,6) the assembled physicists and biologists have arrived at the conclusion that coherent excitation had now been shown to be an intimate feature of biological activity*.

COHERENT EXCITATIONS

Three types of coherent excitation have been proposed:

A. Coherent excitation of a single polar mode.

B. Excitation of a metastable highly polar state; this includes solitons.

C. Vibrations arising from more complex processes and giving rise to limit cycles or Lotka-Volterra oscillations.

All three involve essentially nonlinear processes.

Frequencies of type A processes in membrane vibrations have been estimated in the 10^{11} Hz region; this corresponds to the frequency of millimeter waves. Both, higher and lower frequencies of polar modes in biological molecules, and larger systems, do exist and must be relevant: Such vibrations interact with the abundant cell water and thus will be strongly damped. Spectroscopy of non-excited systems thus shows broad bands only. Coherent excitation, however, as is well known from the theory of lasers, overcomes the damping and leads to sharp lines.

---

*Misrepresentation in this subject is fairly common. Thus in a recent book review in Bioelectromagnetics (December 1982) K.R. Foster seems to be quite unaware that experiments on the influence of millimeter waves on the rate of growth of yeast were undertaken to test the theoretical conjectures.

In the type A model, nonlinear interaction of the polar vibrations with a heat bath under the influence of energy supply at a rate s leads to the transformation of a Planck distribution of excitation at frequency s into a Bose distribution containing a chemical potential μ, i.e. to a replacement of $h\omega$ by $h(\omega-\mu)$. Here increases with increasing s and at a critical value s_o causes Bose condensation, i.e. high excitation of a single mode*.

A metastable highly polar state (B) may arise from the shape dependence of the electrostatic energy of a polarized material (7). Material deformation then must overcome the elastic energy, and the condition for the existence of the state B requires a certain softness which implies the sound velocity c to be below a certain critical value c_o. In extended regions where the sound velocity may vary spatially, those for which $c < c_o$ may exhibit the excitation.

Excitation of type C involves the consequences of A and B and will be discussed below.

Excitation of a metastable state (B) requires a single appropriate energy supply. Excitation of coherent vibrations, however, as mentioned above involves a rate of energy supply s above a critical s_o. When the existing s is just below s_o then small further supply at the rate s_o-s will induce the excitation.

It must also be pointed out that the frequency of the excited vibration need not be exactly reproducible. Assume an ion to be attached to the vibrating region in some cases but not in others. Through electrostatic interaction it may lead to a small shift in frequency through which in course of the nonlinear excitation may be stabilized. Also, nonlinear interaction may induce the frequency shifts which depend on the degree of excitation.

CONSEQUENCES

Coherent excitation of a single polar mode (A) leads to frequency selective long range interaction between systems with equal frequencies; in biological membranes this will be in the 10^{11} Hz region, but higher as well as lower frequencies, originating in other cell structures, will be expected. Such interaction may lead to selective attraction (or repulsion) of systems, to other control of their activities, and to communication within and between cells. It will be noted that electrostatic long range interaction is screened through counter ions.

---

*Extensive theoretical work in this direction has been carried out by Bhaumik et al., Mills and Wu et al.; quoted in (4). Regrettably, very faulty theoretical work has also been published in this field.

Excitation of the metastable polar state (B) may lead to storage and transport of energy. It could represent the active state of enzymes and through the high internal field lead to a reduction of activation energies. It could also be effective in transporting energy from protein to protein. An array of proteins dissolved in a membrane would thus act as the energy transporting network postulated by current thinking (8).

The excitation discussed here may be switched on or off depending on the supply of energy at the appropriate time and place. In general this energy will arise from metabolic processes (or light in case of photosensitivity) and must be regulated by a control system. In the case of excitations A the energy must be supplied to the band of polar modes. According to the theoretical model, the system possesses storage capacity, and a certain time will be required to effect the excitation even if the rate of supply s exceeds the critical one, $s > s_o$. At this stage the system becomes sensitive to interference from the outside. When the relevant band of frequencies is in the microwave, in particular in the millimeter wave region, then low intensity external radiation may cause large biological effects as the switching on of the excitation depends on s in a steplike fashion.

Such effects will be noticeable whenever the externally applied frequency falls into the band of polar modes. When a coherent mode begins to be excited, however, then energy application at its frequency will be more effective. Biological effects dependent on resonance must be expected. The external radiation then acts as a switch. It will be effective at certain periods of the biological development only.

The selective attraction arising in excitation A may play a crucial role in multienzyme processes. Assume enzymes to be spatially organized as discussed by Clegg in (6). Then some of the energy liberated during the enzymatic process may be used to cause excitation A in substrates which then will attract other substrates. This leads to an increase in available substrates and in turn to an increase in the rate of enzymatic processes which decreases the number of substrates. A typical Lotka-Volterra oscillation arises - characteristic of our type C of coherent oscillations. More complex situations may arise leading to limit cycles as discussed in these proceedings by Kaiser. Further consequences are also discussed by Del Guidice.

Such excitations may also be sensitive to external vibrations which might, by relatively weak supply of energy, liberate the stored energy and thus cause large biological effects.

EXPERIMENTS

Existence of three types of coherent excitations: A) a coherent single mode; B) a polar metastable state; C) a periodic activity arising from A and B, arises as the main conclusion of theoretical considerations. Detailed mechanisms might well differ considerably from the particular theoretical models; they might in fact be different in different cases.

Relevant experiments have been discussed in reference (6). They refer to the attraction of nutrients to a cell - similar to the attraction of a substrate as discussed above. The excitation is measured through non thermal excitation of anti-Stokes lines in Raman effect (Drissler et al. in (6). Rowland (6) gives evidence for the mutual attraction of blood cells satisfying the requirements of the theory.

Several contributions are related to biological effects arising from irradiation by very low intensity milimeter waves. Attempts to interpret these in conventional ways, e.g. in terms of the field turning a protein, are quite out of place as they would require (9) intensities of millions W/cm^2 when experiments use 10 mW/cm^2 or less.

Experiments conducted so far do not permit specification of the particular biological process on which electromagnetic fields act. Although recently important evidence in this direction has been obtained, the experiments will have to be continued to establish this evidence. For example, extremely sharp frequency dependence of the influence of millimeter waves on the rate of growth of yeast have been reported by Grundler et al. (6). They find the step-like intensity dependence required by the theory. Apart from a very sharp resonance in the region of 41700 MHz, the frequency dependence exhibits a fine structure of about twenty minima and maxima, spaced at 7-8 MHz. It is tempting to interpret the basic line in terms of a type A excitation with the low frequency arising from processes of type C which require a type A excitation. With very sensitive instrumentation, carried out in an electrically screened laboratory, C.W. Smith and A.H. Jafary have measured very sharp emission in the 7-8 MHz region from single yeast cells. They occur during the period of cell division only, and last only about 3 minutes. The frequencies are not always exactly the same, but this should be interpreted in terms of individuality as discussed in reference 2. We should conclude that the excitation is closely connected with the highly organized processes occurring in cell division.

Long range interaction may also be expected relevant for the control of cell division, important in the cancer problem (4) IVD). Experiments on growing tissue cultures would, therefore, be highly desirable. They do require, however, development of new experimental techniques.

REFERENCES

1. H. Fröhlich, Proc. Natl. Acad. Sci. USA 72:4211 (1975).
2. H. Fröhlich, in: "Theoretical Physics and Biology", M. Marois, ed., North-Holland Publ. Amsterdam, p. 13, (1969).
3. H. Fröhlich, Int. J. Quantum Chem. 2:641 (1968).
4. H. Fröhlich, Advances in Electronics and Electron Phys. 53:85 (1980).
5. F. Keilman and D.B. Kell, Coherent Excitation in Biology, Nature 301:656 (1983).
6. H. Fröhlich and F. Kremer, "Coherent Excitations in Biological Systems" Springer-Verlag, Conference Proceedings - in print (1981).
7. H. Fröhlich, J. Collect. Phenon. 1:101 (1973).
8. D.B. Kell, D.J. Clarke and J.G. Morris, FEMS Microbiology Letters 11:1 (1981).
9. H. Fröhlich, Bioelectromagnetics 3:45 (1982).

MILLIMETER-WAVE AND FAR-INFRARED SPECTROSCOPIC PROPERTIES OF CELLULAR SYSTEMS

K. H. Illinger

Department of Chemistry
Tufts University
Medford, Massachusetts 02155

ABSTRACT

The effect of changes in the general photon distribution on the (infrared-active and Raman-active) vibrational spectra and emissivities of a system of harmonic oscillators is examined. An application is given to the Bose condensation of photons, postulated in the Fröhlich model for _in vivo_ biological systems, and its consequences for their millimeter-wave and far-infrared vibrational spectra and vibrational emissivities. Implications of the analysis for the interpretation of such spectra are discussed.

INTRODUCTION

The spectroscopic properties of _in vivo_ biological systems are determined in part by the (time-invariant) equilibrium system, analyzable in terms of those of ordinary molecular fluids, and in part by non-equilibrium, steady-state (time-variant) subsystems whose analysis may require entirely different constructs. One aspect of such an analysis is the contribution of these subsystems to the vibrational spectrum of cellular systems in the millimeter-wave and far-infrared region. We examine the detailed spectroscopic properties of the model due to Fröhlich (1,2), and the extraction of biophysical information from such spectra.

THE PHOTON DISTRIBUTION FUNCTION

For a molecular system with a set of energy levels m,n,..., interacting with a photon energy density, per unit frequency interval, $\rho(\omega_{mn})$, [erg cm^{-3} sec], at the (angular) transition frequencies, ω_{mn}, the rate equations for the distribution of molecular systems cm^{-3}, $\mathcal{N}_m$, over the states m are given by:

$$\dot{\mathcal{N}}_m = - \mathcal{N}_m \left[\sum_{n \neq m} B_{mn} \rho(\omega_{mn}) \right] + \sum_{n \neq m} \mathcal{N}_n \left[A_{n\to m} + B_{mn} \rho(\omega_{mn}) \right] \qquad m = 1, 2, \ldots \tag{1}$$

Here, the terms B_{mn} are the Einstein coefficients for the photon-induced processes (induced absorption and induced emission) among the states m and n, and $A_{n\to m}$ are the Einstein coefficients for spontaneous emission. In general (3):

$$A_{n\to m} = (\hbar/\pi^2 c^3)\, \omega^3_{mn}\, B_{mn} \tag{2}$$

and:

$$\rho(\omega_{mn}) = (A_{n\to m}/B_{mn})\, n(\omega_{mn}), \tag{3}$$

where $n(\omega_{mn})$ is a (dimensionless) distribution function of photons over frequencies ω_{mn}. The velocity of light in vacuo is given by c, and $\hbar = h/2\pi$, where h is the Planck constant. Thus far, no assumptions have been made concerning the number or the nature of the energy levels of the system, or concerning the form of $n(\omega_{mn})$.

In the case of a two-level system, (m,n), Eq. (1) reduces to:

$$\dot{\mathcal{N}}_m = - \mathcal{N}_m \left[B_{mn} \rho(\omega_{mn}) \right] + \mathcal{N}_n \left[A_{n\to m} + B_{mn} \rho(\omega_{mn}) \right] \tag{4}$$

At steady-state conditions, $\dot{\mathcal{N}}_m=0$, Eq. (4) leads to a simple relation between the molecular-system distribution and the photon distribution:

$$N_m = N_n \{ 1 + [n(\omega_{mn})]^{-1} \} \tag{5}$$

$$(N_m - N_n) = N_m [1 - (N_n / N_m)] = N_m [1 + n(\omega_{mn})]^{-1} \tag{6}$$

Eqs. (5) and (6) express a fundamental property of the two-level system: as the number of photons of frequency ω_{mn} increases, the number of molecular systems in the two levels m and n become equal. If $\rho(\omega_{mn})$ is a Planck distribution for the photons, as expected for equilibrium systems, at a temperature T:

$$n(\omega_{mn}) = n^0(\omega_{mn}) = [e^{\hbar\omega_{mn}/kT} - 1]^{-1}, \tag{7}$$

and there obtains a Boltzmann distribution for the molecular system:

$$(N_m / N_n) = e^{\hbar\omega_{mn}/kT} \tag{8}$$

Here, k is the Boltzmann constant.

Although for the multilevel system appropriate to this treatment, one must strictly use the coupled equations, Eq. (1), the behavior of the two-level system serves as a guide, via Eq. (5), in exploring the consequences of general photon distributions, $n(\omega_{mn})$, other than the Planck distribution. In particular, the steady-state boson distribution arising in the Fröhlich vibrational model (1,2) for *in vivo* cellular systems may initially be examined in this fashion. For first-order and second-order exchange processes between the non-equilibrium subsystem and the remainder of the system, $n(\omega_{mn})$ takes on the form (2):

$$n(\omega_k)_1 = [e^{\hbar\omega_k/kT} - 1]^{-1} [1 + S_k/\phi] \tag{9}$$

and:

$$n(\omega_k) = \{1 + a_k(1 - e^{-\mu/kT})\} [e^{[(\hbar\omega_k - \mu)/kT]} - 1]^{-1} \tag{10}$$

Here, k denotes the k'th vibrational mode, with a frequency $\omega_{mn}=\omega_k$, $(m=v_k,\ n=v_k+1)$, and $a_k=(S_k/S)(\phi/\chi)$, $S=\sum_k S_k$. For $S_k=\mu=0$, a Planck distribution results from Eqs. (9) and (10). The pumped system, $S_k>0$, in the absence of non-linear terms in Eq. (10), i.e. $\mu=0$, is represented by Eq. (9), a Planck-type distribution modified by the multiplicative factor $(1+S_k/\phi)$. In the presence of non-linear exchange between the subsystem and the remainder of the system, $\mu>0$, and for $S>S_0$, the complex situation represented by Eq. (10), one may write for the lowest-frequency mode ω_1, which exhibits the Bose-type condensation of photons:

$$\lim_{\substack{(\hbar\omega_1-\mu)/kT \ll 1 \\ S>S_0}} n(\omega_1) = \left(\frac{kT}{\hbar\omega_1-\mu}\right)\left\{1+a_1\left(1-e^{-\mu/kT}\right)\right\} < N(S,T) \tag{11}$$

Here, $N(S,T)=\sum_k n(\omega_k)$ is the total number of photons in all modes k. As μ approaches closely to $\hbar\omega_1$, the number of photons in the lowest mode increases inversely with the term $(\hbar\omega_1-\mu)$. In the limit, $(\hbar\omega_1/kT)\ll 1$, appropriate if ω_1 lies well below $\sim$ 200 cm^{-1} ($\sim$6000 GHz) at a temperature of the equilibrium system at $\sim$ 300°K, Eq. (11) simplifies further to:

$$n(\omega_1) \cong \left(\frac{kT+a_1\mu}{\hbar\omega_1-\mu}\right) < N(S,T) \tag{12}$$

The salient difference between Eq. (12) and the low-frequency limit, $(\hbar\omega_1/kT)\ll 1$ of the Planck distribution, Eq. (7), $n^{\circ}(\omega_1) = (kT/\hbar\omega_1)$, lies in the magnitude of μ, $\hbar\omega_1>\mu\geq 0$, the measure of the non-linear interaction terms in the Fröhlich vibrational model. As $(\hbar\omega_1-\mu)/kT\rightarrow 0$, $n(\omega_1)$ rapidly increases with decreasing $(\hbar\omega_1-\mu)/kT$. As a result, for the molecular system bridged by the photons ω_1, causing the transitions $v_1\leftrightarrow v_1+1$, one has from Eqs. (5) and (12), in the two-level approximation.

$$N_{v_1} = N_{v_1+1}\left[1+\left(\frac{\hbar\omega_1-\mu}{kT+a_1\mu}\right)\right] \tag{13a}$$

and:

$$(N_{v_1} - N_{v_1+1}) = N_{v_1}\left[1 + \left(\frac{\hbar\omega_1 - \mu}{kT + a_1\mu}\right)^{-1}\right] \tag{13b}$$

For the vibrational modes $k \neq 1$, $\omega_k > \omega_1$, in the Fröhlich model, Eq. (10) applies. For those modes for which $\omega_k \gg \omega_1 \cong (\mu/\hbar)$, this implies a Planck distribution, modified by a multiplicative factor, $(1+a_k\mu/kT)$, concomitant with the Bose condensation for the lowest-frequency mode, ω_1; namely:

$$\lim_{\substack{(\hbar\omega_1-\mu)/kT \ll 1 \\ S > S_0}} n(\omega_k) = \left\{1 + \frac{a_k\mu}{kT}\right\}\left[e^{\hbar\omega_k/kT} - 1\right]^{-1} \tag{14}$$

Eq. (14) is consistent, via Eq. (5), with a molecular-systems distribution:

$$N_{v_k} = N_{v_k+1}\left\{1 + \frac{a_k\mu}{kT}\right\}^{-1}\left[e^{\hbar\omega_k/kT} + \frac{a_k\mu}{kT}\right] \tag{15}$$

To the extent to which $0 < a_k\mu \ll kT$, the high-frequency modes are representable by a modified Boltzmann distribution, in marked contrast to the prediction of Eq. (13) for the lowest mode, k=1.

The foregoing considerations have been based, subsequent to Eq. (4), on the characteristics of the two-level system. Actual molecular systems, with rare exceptions, are multilevel systems, as are those appropriate to the Fröhlich vibrational model. In spite of this, the assumption usually made in the elementary analysis of spectra, is that the behavior of multilevel systems is predictable by a concatenation of (uncoupled) two-level systems, tractable by Eq. (4). It is of interest to relax this restriction in the present application. If the set of harmonic-oscillator systems in the Fröhlich model is mechanically uncoupled (1), the problem reduces to that of the separate multilevel systems, $v_k=0,1,2, \ldots$ Ultimately, for characterization of the spectroscopic properties of this multilevel, multimode ($k=1,2, \ldots z$) system, we shall require the total number of molecular systems cm^{-3} in a state v_k, irrespective of their quantum numbers in the other modes, v_j; namely:

$$\mathcal{N}_{v_k} = \sum_{j\neq k} \left\{ \sum_{v_j=0}^{\infty} \mathcal{N}(v_1, v_2, \ldots v_j, v_k, \ldots v_z) \right\} , \tag{16}$$

where $\mathcal{N}(v_1,v_2,..v_j,..v_k,..v_z)$ is the molecular-system distribution function in a completely specified vibrational state. Initially, it is necessary to examine the separate multilevel systems, v_k.

For a non-degenerate harmonic oscillator, with vibrational states $v=v_k=0,1,2,..$, Eq. (1), in the harmonic-oscillator, linear dipole-moment approximation reduces to:

$$\left\{ \begin{aligned} \dot{\mathcal{N}}_v = & - \mathcal{N}_v \left[B_{v\to v+1}\, \rho(\omega_{v,v+1}) + B_{v\to v-1}\, \rho(\omega_{v,v-1}) + A_{v\to v-1} \right] \\ & + \mathcal{N}_{v+1} \left[A_{v+1\to v} + B_{v+1\to v}\, \rho(\omega_{v+1,v}) \right] \\ & + \mathcal{N}_{v-1}\, B_{v-1\to v}\, \rho(\omega_{v-1,v}) \\ v = & \; 0,1,2,\ldots \end{aligned} \right. \tag{17}$$

Although this approximation, limiting photon-induced transitions to those between adjacent states, v↔v±1, is not rigorously correct for any real system, its employment in inquiring about the overall population statistics will not be unduly restrictive. Making use of the relations expressed by Eq. (4),

$$\begin{aligned} \rho(\omega_{vv'}) &= (A_{v'\to v} / B_{vv'})\, n(\omega_k) = (\hbar/\pi^2 c^3)\, \omega_k^3\, n(\omega_k) \\ A_{v\to v-1} &= v\, A_{10} \\ \omega_{vv'} &= \omega_k , \end{aligned} \tag{18}$$

where A_{10} is the Einstein coefficient for spontaneous emission between the states 1→0 and $n(\omega_k)$ is the generalized photon distribution function, one obtains, after some manipulation:

$$\begin{cases} \dot{N}_v = A_{10}\left\{ n(\omega_k)\, v\, N_{v-1} - \left[n(\omega_k)\,(2v+1) + v \right] N_v \right. \\ \qquad\qquad \left. + \left[1 + n(\omega_k) \right] (v+1)\, N_{v+1} \right\} \\ v = 0,1,2,\ldots \end{cases} \tag{19}$$

For steady-state conditions, $\dot{N}_v=\dot{n}(\omega_k)=0$; in consequence:

$$N_v = \left[(2v+1)\, n(\omega_k) + v \right]^{-1} \left\{ (v+1) \left[1 + n(\omega_k) \right] N_{v+1} + v\, n(\omega_k)\, N'_{v-1} \right\}$$

$$v = 0,1,2,\ldots \tag{20}$$

Because of the highly restrictive nature of the selection rules implicit in Eq. (17), one finds for Eq. (20) a solution for the molecular-systems distribution N_v closely resembling that in the two-level approximation, Eq. (5); namely:

$$N'_v = N_{v+j} \left\{ 1 + \left[n(\omega_k) \right]^{-1} \right\}^{j}$$

$$|j| = 1,2,3,\ldots \tag{21}$$

For a constant total number of molecular systems cm^{-3}, N, at steady-state conditions, $\dot{N}=0$,

$$N = \sum_{v=0}^{\infty} N_v = N_0 \sum_{v=0}^{\infty} \left\{ 1 + \left[n(\omega_k) \right]^{-1} \right\}^{-v} = N'_0 \left[1 + n(\omega_k) \right] \tag{22}$$

The vibrational partition function, θ_k, for the generalized photon distribution $n(\omega_k)$, and for a mode degeneracy, $g_k=1,2,3$, is given by:

$$\theta_k = N/N_0 = \sum_{v_k=0}^{\infty} g(v_k)\, N_{v_k} = \left[1 + n(\omega_k)\right]^{g_k}, \tag{23}$$

where $g(v_k)$ is the degeneracy of the level v_k (5). Eqs. (21) and (23) permit one to write a general expression for the total number of molecular systems cm^{-3} in the level v_k of the multimode system from Eq. (16),

$$N_{v_k} = \frac{N}{\theta_k}\, g(v_k) \left\{1 + \left[n(\omega_k)\right]^{-1}\right\}^{-v_k} \tag{24}$$

For the Bose-condensation distribution, Eq. (12). Eq. (23) becomes:

$$\theta_1 \cong \left[1 + \left(\frac{kT + a_1\mu}{\hbar\omega_1 - \mu}\right)\right]^{g_1} = \left(\frac{kT_{eff}(\omega_1)}{\hbar\omega_1}\right)^{g_1} \tag{25}$$

In the last term of Eq. (25), T_{eff} may be interpreted as a (fictitious) effective temperature for the non-Boltzmann distribution on the molecular subsystem in mode ω_1. For $(\hbar\omega_1-\mu)\rightarrow 0$, $T_{eff}(\omega_1) > T$, the (actual) temperature of the remainder of the molecular system. Eq. (16), applied to the lowest-frequency mode ω_1, becomes, via Eqs. (12) and (24):

$$N_{v_1} \cong \frac{N}{\theta_1}\, g(v_1) \left[1 + \left(\frac{\hbar\omega_1 - \mu}{kT + a_1\mu}\right)\right]^{-v_1} \tag{26}$$

If the approximation $(\hbar\omega_1/kT) \ll 1$ is not valid, Eq. (11) must be used to develop equations analogous to Eqs. (23) and (24). For $n(\omega_k)$ a Planck distribution, Eq. (7), Eq. (20) takes on the form (6):

$$\left\{\begin{array}{l} \mathcal{N}_v = [(v+1)e^{-x_k} + v]^{-1}\,[(v+1)\,\mathcal{N}_{v+1} + v\,e^{-x_k}\,\mathcal{N}_{v-1}] \\ v = 0,1,2,\ldots \end{array}\right. \tag{27}$$

and Eq. (21), with x_k defined as $(\hbar\omega_k/kT)$, reduces to:

$$\mathcal{N}_v = \mathcal{N}_{v+j}\;e^{jx_k} \tag{28}$$

Onset of the Bose condensation in the nonequilibrium subsystem(s) leads to conversion of the term $n(\omega_1)$ from a Planck distribution to a distribution of the form given by Eqs. (10), (11) and (12). Excitation of the Bose condensation, in a coherent mode in cellular systems, is postulated (7-11) to have fundamental implications for cellular biophysics. We wish here to examine further its implications for the detailed spectroscopic properties of the nonequilibrium subsystem(s): (i) the contribution to the total attenuation function $\alpha(\omega)$, (ii) Raman-spectroscopic transitions, and (iii) the contribution to the emissivity.

THE ATTENUATION FUNCTION

The total subsystem attenuation function $\alpha(\Delta v_k=1)$ [neper cm^{-1}] due to all the fundamental vibrational transitions, $v_k \leftrightarrow v_k \pm 1$, for a nondegenerate harmonic oscillator, in the strict harmonic-oscillator - linear dipole-moment approximation, is given by (12,5):

$$\begin{aligned} \alpha(\Delta v_k=1) &= \sum_{v_k=0}^{\infty} \alpha(v_k, v_k+1) \\ &= \frac{4\pi}{3\hbar c}\,\omega_k\,|M_{01}|_k^2 \sum_{v_k=0}^{\infty} (\mathcal{N}_{v_k} - \mathcal{N}_{v_k+1})(v_k+1) \end{aligned} \tag{29}$$

Here, $|M_{01}|_k$ is the dipole-moment matrix element for the transition $0 \leftrightarrow 1$. In order to assess the effect of the redistribution of molecular systems among the states v_k, due to changes in the general photon distribution function $n=n(\omega_k)$, on $\alpha(\Delta v_k=1)$, we require evaluation of the sum on the rhs of Eq. (29). Employing Eqs. (23) and (24) for this multilevel, multimode system, one obtains:

$$\sum_{v_k=0}^{\infty} (\mathcal{N}_{v_k} - \mathcal{N}_{v_k+1})(v_k+1) = \frac{\mathcal{N}}{\Theta_k}(1+n)^{-1} \sum_{v_k=0}^{\infty} [1+n^{-1}]^{-v_k}(v_k+1)$$

$$= \mathcal{N} \tag{30}$$

since $\Theta_k = (1+n)$, and the sum in the second term of Eq. (30) is $(1+n)^2$. As a result, from Eqs. (29) and (30):

$$\alpha(\Delta v_k = 1) = \frac{4\pi}{3\hbar c} \mathcal{N} \omega_k |M_{01}|_k^2 \neq f[n(\omega_k)] \tag{31}$$

Thus, in the harmonic-oscillator linear dipole-moment approximation, the total attenuation function, which has contributions from absorptive and from emissive photon-induced processes, is unaffected by alterations of the general photon distribution $n(\omega_k)$. Extension of the arguments embedded in Eqs. (29) and (30) to degenerate harmonic-oscillator systems, $g_k = 1,2,3$, does not alter this conclusion. The invariance of $\alpha(\Delta v_k = 1)$ is a consequence of the highly restrictive selection rules embodied in Eqs. (20) and (29), the fact that all the allowed (fundamental) transitions occur at the same frequency, ω_k, and the dependence of the transition probabilities on the vibrational state v_k. For the multilevel system appropriate to the Fröhlich model, while $n(\omega_k) \gg 1$ reduces the population differences in Eq. (29), the concomitant increase in the transition probabilities $v_k \leftrightarrow v_k \pm 1$, as molecular systems are pumped into excited vibrational states, exactly balances the effect of this reduction, in the sum of terms $\alpha(v_k, v_k+1)$, and renders $\alpha(\Delta v_k = 1)$ invariant to changes in $n(\omega_k)$. This is in contrast with the properties of a two-level system, which exhibits saturation, with $\alpha(v_k, v_k+1) \rightarrow 0$ as $n(\omega_k) \gg 1$. Thus, to first order, the onset of the Bose condensation in the Fröhlich vibrational model will not affect the contribution to the total attenuation function of the fundamental vibrations, in any of the modes, k, of the nonequilibrium subsystem(s).

While the strict harmonic-oscillator linear dipole-moment approximation limits all photon-induced vibrational transitions to fundamental vibrations, non-linearity in the dipole-moment function, even if the vibrations are mechanically harmonic, is associated with the existence of overtone vibrations, $v_k \leftrightarrow v_k + j_k$,

$j_k=2,3,\ldots$ Contributions to $\alpha(\omega)$ from overtone vibrations may be formally represented by:

$$\alpha(\Delta v_k = j_k) = \sum_{v_k=0}^{\infty} \alpha(v_k, v_k+j_k)$$

$$= \frac{4\pi}{3\hbar c} \sum_{v_k=0}^{\infty} \omega(v_k, v_k+j_k)\,(N_{v_k} - N_{v_k+j_k})\,|M_{v_k, v_k+j_k}|^2 \qquad (32)$$

Here, the overtone frequencies $\omega(v_k, v_k+j_k)$ are approximately given by $\omega_k j_k$, in the limit of small anharmonicities. The dipole moment matrix elements $|M_{v_k, v_k+j_k}|$ have been estimated in terms of a simplified model (13), and are given, to the same approximation, neglecting mechanical anharmonicity, by:

$$|M_{v_k, v_k+j_k}|^2 = |M_{0, j_k}|^2\,(v_k+j_k)!\,(v_k!)^{-1}\,(j_k!)^{-1} \qquad (33)$$

They increase, for a given j_k, by a factor $[(v_k+j_k)!/(v_k!)$. Exact evaluation of the sum on the rhs of Eq. (32) would require recomputation of the population distribution, Eq. (1), including the overtone transitions $m=v_k \leftrightarrow n=(v_k+j_k)$ disallowed in Eqs. (19) and (20), in terms of the matrix elements $|M_{0,j_k}|$, which are not known for the complex molecular systems being considered here. The appropriate equations, generalizations of Eqs. (19) and (20) would further contain all the photon distribution-function terms in the vicinity of each of the frequencies $\omega_k j_k$, $j_k=2,3,\ldots$ Qualitatively, arguments similar to those given for the fundamental transitions should apply to overtones, and $\alpha(\Delta v_k=j_k)$ is expected to be insensitive to alterations in $n(\omega_k j_k)$. Neglect of the effect on the population distribution due to the photon distribution at overtone frequencies is justifiable _a fortiori_ in situations, as in the Fröhlich vibrational model, where the fundamental-frequency term, $n(\omega_1)$, is expected to outweigh all others, i.e., $n(\omega_1) \gg n(\omega_k j_k)$, $\omega_k j_k \neq \omega_1$.

Experimental determination of the fundamental and overtone transitions, $\alpha(\Delta v_k=j_k)$, $k=1,2,\ldots$, would provide the frequencies ω_k, if proper assignments can be made to the vibrational modes of the subsystem(s). However, to the approximation investigated here, attenuation measurements provide no information about the other parameters of the Fröhlich model, nor do they monitor the onset and/or collapse of the Bose condensation. While accurate measurements of the attenuation function of the dissipative subsystem(s) are impeded, particularly in the millimeter-wave and far-infrared region, by the dominance of absorption by the remainder of the molecular system (especially $H_2O(1)$), millimeter-wave absorption measurements (14) are in accord with the predictions of Eq. (31).

RAMAN INTENSITIES

The invariance of the attenuation function with alterations in the photon distribution function, as the subsystem is driven from equilibrium and experiences the attendant redistribution over states of its molecular systems, arises from the specific form of the detailed balance of the interlaced absorption and emission processes among the states of the harmonic-oscillator systems. This invariance is not expected for processes which depend on the separate sets of upward or downward molecular transitions, $v_k \rightarrow v_k + j_k$ and $v_k + j_k \rightarrow v_k$, respectively. Such processes are inelastic scattering in the vibrational Raman spectrum and the vibrational emissivity. The absolute intensities of a set of downward (anti-Stokes) and upward (Stokes) transitions, $v_k + j_k \rightarrow v_k$ and $v_k \rightarrow v_k + j_k$, are given by (15,5):

$$I[\omega_0 + \omega(v_k, v_k + j_k)]$$

$$= \frac{8\pi}{3c^4} I_0 [\omega_0 + \omega(v_k, v_k + j_k)]^4 \mathcal{N}_{v_k + j_k} |P_{v_k, v_k + j_k}|^2 \tag{34a}$$

and:

$$I[\omega_0 - \omega(v_k, v_k + j_k)]$$

$$= \frac{8\pi}{3c^4} I_0 [\omega_0 - \omega(v_k, v_k + j_k)]^4 \mathcal{N}_{v_k} |P_{v_k, v_k + j_k}|^2 \tag{34b}$$

Here, ω_o is the frequency and I_o the intensity of the exciting line, and $|P_{v_k,v_k+j_k}|$ is the matrix element for the induced dipole moment. If it is assumed, as in the above, that deviations from mechanical anharmonicity are negligible, all the frequencies $[\omega_o \pm \omega(v_k,v_k+j_k)] \cong [\omega_o \pm \omega_k j_k]$, and the total Raman intensities become:

$$I(\omega_0+\omega_k j_k) = \frac{8\pi}{3c^4} I_0 (\omega_0+\omega_k j_k)^4 \sum_{v_k=0}^{\infty} \mathcal{N}_{v_k+j_k} |P_{v_k,v_k+j_k}|^2 \tag{35a}$$

and:

$$I(\omega_0-\omega_k j_k) = \frac{8\pi}{3c^4} I_0 (\omega_0-\omega_k j_k)^4 \sum_{v_k=0}^{\infty} \mathcal{N}_{v_k} |P_{v_k,v_k+j_k}|^2 \tag{35b}$$

As in the case of the attenuation function, the population terms $\mathcal{N}_{v_k}$ and $\mathcal{N}_{v_k+j_k}$ are given exactly by an equation of the type of Eq. (1). If we assume, as in the above, that only fundamental photon-induced transitions are important in the calculation of the $\mathcal{N}_{v_k}$, Eq. (24) may be employed, with $g_k=1$, in conjunction with Eq. (35), to give for the total Raman intensities of the anti-Stokes and Stokes lines, respectively:

$$\begin{aligned} I(\omega_0+\omega_k j_k) &= \frac{8\pi\mathcal{N}}{3c^4\theta_k} I_0 (\omega_0+\omega_k j_k)^4 \{1+[n(\omega_k)]^{-1}\}^{-j_k} \\ &\cdot \sum_{v_k=0}^{\infty} \{1+[n(\omega_k)]^{-1}\}^{-v_k} |P_{v_k,v_k+j_k}|^2 \end{aligned} \tag{36a}$$

and:

$$I(\omega_0-\omega_k j_k)$$

$$= \frac{8\pi N}{3c^4\theta_k} I_0(\omega_0-\omega_k j_k)^4 \sum_{v_k=0}^{\infty} \left\{1+[n(\omega_k)]^{-1}\right\}^{-v_k} \left|P_{v_k,v_k+j_k}\right|^2 \tag{36b}$$

The anti-Stokes to Stokes Raman-intensity ratio is seen to be independent of the form of the matrix elements $\left|P_{v_k,v_k+j_k}\right|$, and to be a function of the photon distribution $n(\omega_k)$; namely:

$$\frac{I(\omega_0+\omega_k j_k)}{I(\omega_0-\omega_k j_k)} = b_{kj}^4 \left\{1+[n(\omega_k)]^{-1}\right\}^{-j_k} \tag{37}$$

As in the case of Eq. (31), Eq. (37) holds also for degenerate vibrations, $g_k=1,2,3$. In the limit, $n(\omega_k) \gg 1$, the Raman-intensity ratio approaches unity, since, in a typical Raman experiment, $\omega_o \cong 20000\ \mathrm{cm}^{-1}$, $\omega_o \gg \omega_k j_k$, and $b_{jk} \cong 1$, where: $b_{jk}=(\omega_o+\omega_k j_k)/(\omega_o-\omega_k j_k)$.

For the lowest-frequency vibration in the Fröhlich vibrational model, Eqs. (11) and (12), one obtains from Eq. (37):

$$\frac{I(\omega_0+\omega_1 j_1)}{I(\omega_0-\omega_1 j_1)} = b_{1j}^4 \left\{1+\left(\frac{\hbar\omega_1-\mu}{kT}\right)\left[1+a_1\left(1-e^{-\mu/kT}\right)\right]^{-1}\right\}^{-j_1}$$

$$\cong b_{1j}^4 \left\{1+\left(\frac{\hbar\omega_1-\mu}{kT+a_1\mu}\right)\right\}^{-j_1}$$

$$\cong b_{1j}^4 \left\{1+\left(\frac{\hbar\omega_1-\mu}{kT}\right)\right\}^{-j_1} \tag{38}$$

In Eq. (38), the first approximation (line 2) assumes, $(\hbar\omega_1/kT) \ll 1$; the second (line 3), $(a_1\mu/kT) \ll 1$.

For $\mu=0$, Eqs. (10) and (37) reduce to a Boltzmann-type intensity ratio:

$$I(\omega_0+\omega_k j_k)/I(\omega_0-\omega_k j_k) = b_{kj}^4\, e^{-j_k x_k} \tag{39}$$

For $k\neq 1$, $\omega_k \gg \omega_1 \cong (\mu/\hbar)$, the Raman intensity ratio may be computed employing Eqs. (14) and (37):

$$\frac{I(\omega_0+\omega_k j_k)}{I(\omega_0-\omega_k j_k)} = b_{kj}^4\left(1+\frac{a_k\mu}{kT}\right)\left(e^{x_k}+\frac{a_k\mu}{kT}\right)^{-j_k} \tag{40}$$

If the inequality cited immediately above does not hold, the Raman-intensity ratio for a mode k is given by Eqs. (10) and (37).

Eqs. (37), (38) and (40) provide an experimental avenue toward the evaluation of the parameters μ and a_k in the Fröhlich vibrational model, if the frequencies ω_k, and hence the j_k's, can be assigned. While the Raman-intensity ratios for the modes $k\neq 1$ are not expected to change dramatically as $(\hbar\omega_1-\mu)/kT \rightarrow 0$, the Raman-intensity ratios for the lowest-frequency mode, k=1, internally pumped by the Bose condensation of photons in ω_1, should show a major alteration in this limit. This has been claimed to occur in the low-frequency (< 200 cm^{-1}) Raman spectrum during certain temporal phases of cell growth in E. coli (16).

THE VIBRATIONAL EMISSIVITY

From Eqs. (2) and (3), the photon energy density, per unit frequency interval, at a fundamental frequency ω_k of the subsystem is given by:

$$\rho(\omega_k) = (\hbar/\pi^3 c^3)\,\omega_k\, n(\omega_k) \tag{41}$$

The emission intensity (or radiancy), $R(\omega_k)$, [erg cm^{-2} sec^{-1}], i.e. the photon energy emitted across a unit surface area, per unit time, into a solid angle of 2π steradians, is related to $\rho(\omega_k)$ by (17):

$$R(\omega_k) = (c/4)\,\rho(\omega_k) = (\hbar/4\pi^2 c^2)\,\omega_k^3\, n(\omega_k) \tag{42}$$

The emissivity, $\varepsilon(\omega_k)$, defined as the ratio of the actual emission intensity to that of an ideal blackbody at the same temperature, then is, at the surface of the emitting subsystem:

$$\varepsilon(\omega_k) = [R(\omega_k) / R^{o}(\omega_k)] = [n(\omega_k) / n^{o}(\omega_k)] \tag{43}$$

Here, $R^{o}(\omega_k)$ is the radiancy of an ideal blackbody, and $n^{o}(\omega_k)$ is the Planck distribution, Eq. (17).

For the Bose-condensation conditions appropriate to Eqs, (11) and (12), the vibrational emissivity, $\varepsilon(\omega_1)$, at the frequency ω_1 in situ of the subsystem, is given, to the same approximation as that pertinent to Eq. (38), by:

$$\begin{aligned}
\varepsilon(\omega_1) &= \left(\frac{kT}{\hbar\omega_1 - \mu}\right)\left(e^{x_1} - 1\right)\left[1 + a_1\left(1 - e^{-\mu/kT}\right)\right] \qquad \gg 1 \\
&\cong \left(\frac{kT + a_1\mu}{\hbar\omega_1 - \mu}\right) x_1 \\
&\cong \left(\frac{\hbar\omega_1}{\hbar\omega_1 - \mu}\right)
\end{aligned} \tag{44}$$

In contrast, the vibrational emissivities at the frequencies ω_k, $k \neq 1$, are, according to Eqs. (10), (14) and (23):

$$\begin{aligned}
\varepsilon(\omega_k) &= \left[1 + a_k\left(1 - e^{-\mu/kT}\right)\right]\left\{e^{[(\hbar\omega_k - \mu)/kT]} - 1\right\}^{-1}\left(e^{x_k} - 1\right) \\
&\cong \left[1 + a_k\left(1 - e^{-\mu/kT}\right)\right] \\
&\cong \left(1 + \frac{a_k\mu}{kT}\right) \\
&\cong 1
\end{aligned} \tag{45}$$

In Eq. (45), the first approximation (line 2) assumes, $\omega_k \gg \omega_1 \cong (\hbar/\mu)$; the second (line 3), $x_k \ll 1$; the third (line 4), $(a_k\mu/kT) \ll 1$.

For $(\hbar\omega_1 - \mu)/kT \ll 1$, the emissivity *in situ* of the nonequilibrium subsystem, $\varepsilon(\omega_1)$, has a maximum at ω_1, directly proportional to the ratio $[\hbar\omega_1/(\hbar\omega_1 - \mu)]$. Experimental measurement of the radiancy $R(\omega_1)$, which provides the emissivity term $\varepsilon(\omega_1)$, in principle would determine both ω_1 and the magnitude of μ, via Eq. (44), and analogous determination of the quantities $R(\omega_k)$ calibrates $(a_k\mu/kT)$. However, several circumstances conspire to impede measurements of the emissivity contributed to by the subsystem(s), in a typical biological context. The total concentration, and hence the surface area of the subsystem(s) may typically be very small, in comparison with the total system, which will contribute roughly a blackbody background. Furthermore, it is not known what are the widths of the emission lines, and the experimental variable, $R(\omega_k)$, Eq. (42), is more generally given by:

$$R(\omega_k) = \varepsilon(\omega_k)\, R^0(\omega_k) = (\hbar/4\pi^2c^2)\, \omega_k^3\, n(\omega_k)\, D_k(\omega) \tag{46}$$

Here, $D_k(\omega)$ is an emission bandshape function. The greater the halfwidth of $D_1(\omega)$, the more difficult is it experimentally, for a given signal-to-noise ratio, to discern the features of $\varepsilon(\omega_1) > 1$, in the presence of the background contribution due to the remainder of the system. Similar considerations concerning the concentration and the line-width apply, of course, also to the subsystem contribution to the attenuation function and the individual Raman intensities, Eq. (36). Finally, for ω_k in the range 1-10^2 cm^{-1}, the attenuation function due to the total system is typically very large, of the order of 10-10^2 neper/cm, and the emission intensity of subsystems in the interior of a cellular system may be strongly attenuated in its passage through the total system. Strong attenuation of the Raman-scattering lines at $(\omega_0 \pm \omega_k j_k)$ is not expected. The details of this effect depend upon the spatial distribution and local concentration of the emitting system, and the dielectric properties, at the same frequency, of the total system. For a homogeneous distribution of subsystems and a homogeneous total-system attenuation function, $\alpha(\omega_k)$, the emissivity at an optical depth z is (17):

$$\varepsilon(\omega_k, z) = \left[1 - e^{-\alpha(\omega_k) z}\right] \left[R(\omega_k)/R^0(\omega_k)\right] \tag{47}$$

For a single layer, at z, of subsystems emitting at an optical depth z,

$$\varepsilon(\omega_k, z) = e^{-\alpha(\omega_k)z}\left[R(\omega_k)/R^0(\omega_k)\right] \tag{48}$$

Actual cellular systems are expected to be considerably more complex than these idealized extreme cases.

DISCUSSION

Of the three spectroscopic observables of the non-equilibrium subsystem(s), the emissivity, $\varepsilon(\omega_k j_k)$, the Raman-intensity ratio, $[I(\omega_o+\omega_k j_k)/I(\omega_o-\omega_k j_k)]$, and the attenuation function, $\alpha(\Delta v_k=j_k)$, the first would provide the most direct experimental evidence of the quantitative characterization of the parameters of the Fröhlich vibrational model: the frequency of the Bose-condensation vibrational mode, ω_1, the measure of the magnitude of the nonlinear terms, μ, and the parameter a_1. The terms $\alpha(\Delta v_k=j_k)$ are invariant with $n(\omega_k)$, hence provide no information concerning the onset of the Bose condensation, and preclude facile assignment of the frequencies ω_k. The Raman-intensity ratios, while dependent on the form of $n(\omega_k)$, are insensitive to the value of j_k, in the limit $n(\omega_k) \gg 1$, as seen from Eq. (37). In the strict harmonic-oscillator linear dipole-moment approximation, only the fundamental terms $\alpha(\Delta v_k=1)$ occur, provided that the matrix elements in Eq. (31) are non-vanishing. Although the general theory of the (Raman) induced dipole-moment matrix elements $|P_{v_k,v_k+j_k}|$ (18) is computationally intractable, the polarizability theory of the Raman effect (19) is consistent with the same selection rule, $\Delta v_k \underline{+} 1$, if a linear polarizability function is assumed. In fact, vibrational Raman spectra rarely show overtone or combination bands, unless these are made active through Fermi resonance with another vibration of the same symmetry species. This is in contrast to the not infrequent occurence of these transitions in the infrared-active vibrational spectrum. However, this empirical rule, which tends to imply stronger linearity in the polarizability function (with respect to the vibrational displacement coordinate) than in the dipole-moment function, applies principally to group vibrations of a molecular system. The vibrational distortions implicated in the model encompass very large molecular systems, and their properties are not as well characterized as the normal modes of a simple molecular system. Strict transfer of these findings for molecular systems would suggest a single Raman-active fundamental transition, provided the matrix element $|P_{0,1}|_k$, Eq. (36) with $j_k=1$, is non-vanishing, for each mode k. Observation of the lowest-frequency mode, its dependence on (time-variant) cellular activity and its attendant alteration in $n(\omega_1)$, would then, in principle, assign the frequencies ω_k. Apart from the question as to whether these selection rules apply strictly to the complex systems of interest here, this procedure is impeded by the fact that, in the limit of low frequencies, $x_k \ll 1$, the Raman-intensity ratio provides no basis for differentiation for a given experimental error in intensity measurement, between the (equilibrium) Planck distribution $n^o(\omega_k)$ and a non-equilibrium distribution with $n(\omega_k) \gg 1$. In consequence, characterization of the Bose condensation by the Raman-intensity ratio is the more reliable the less stringent the inequality, $x_k \ll 1$. Thus, the experimental demonstration of the Bose condensation via the Raman-intensity ratio is as yet of a preliminary nature. In particular, it is not known what relation the reported spectra (16) bear, if any, to the highly frequency-selective biological response functions (20-22) that have been reported in the millimeter-wave region.

Certain of the experimental and interpretational problems attaching to these investigations have been discussed in some detail elsewhere (23-25). Raman spectra of in vivo systems (26), demonstrating the appearance of Raman-active transitions in photosynthetic bacteria, have been described, attributing these transitions to the formation of supramolecular systems active in the photosynthetic process. General non-equilibrium statistical-thermodynamic criteria can be adduced (27) to obtain and analyze the boson (photon and phonon) distribution function, of the form of Eq. (10), which accompanies the onset ($\mu > 0$) of non-equilibrium thermodynamic processes in the dissipative molecular subsystem(s).

A single emissivity term, $\varepsilon(\omega_1)$, Eq. (44), is predicted to exhibit behavior saliently different from all the others, $\varepsilon(\omega_k)$, $k \neq 1$, each of these being essentially blackbody terms, $\varepsilon(\omega) \cong 1$. In consequence, $\varepsilon(\omega_1) \gg 1$ directly defines the lowest-mode frequency, ω_1, and the magnitude of μ, the crucial parameter in the Fröhlich model. For the other modes, $k \neq 1$, $\varepsilon(\omega_k)$ in principle provides information concerning the products $a_k \mu$. In spite of the experimental problems concerning the determination of the emissivity of the nonequilibrium subsystem(s) discussed in the above, this experimental variable would appear to provide the most direct route to the quantification of the model. In general, direct spectroscopic determinations of the millimeter-wave and far-infrared properties of in vivo cellular systems possess the not inconsiderable advantage of providing biophysical information at the systems level, while global biological endpoints are intrinsically integrative effects, not currently susceptible of direct analysis.

REFERENCES

1. H. Fröhlich, Long-range Coherence in Biological Systems, *Revista del Nuovo Cim.* 7:399 (1977)

2. H. Fröhlich, in *Advances in Electronics and Electron Physics*, *53*: 85, L. Marton and C. Marton, Ed. Academic Press, New York, (1980).

3. J. C. Slater, " Quantum theory of Atomic Structure ", McGraw-Hill, New York (1960), Vol. I, Chap. 6.

4. L. Pauling and E.B. Wilson, " Introduction to Quantum Mechanics " McGraw-Hill, New York (1935), p. 77 ff.

5. K.H Illinger and D.E. Freeman, Temperature Dependence of Infrared and Raman Intensities , *J. Molec. Spectroscopy* 9:191 (1962).

6. R.J. Rubin and K.E. Shuler, On the Relaxation of Vibrational Equilibrium Distributions. III. The Effect of Radiative Transitions on the Relaxation Behavior , *J. Chem. Phys.* 26:137 (1957).

7. H. Fröhlich, Long-range Coherence and Energy Storage in Biological Systems, *Intern. J. Quant. Chem.* 2:641 (1968).

8. H. Fröhlich, Bose Condensation of Strongly Excited Longitudinal Electric Modes, *Phys. Lett.* 26A:402 (1968).

9. H. Fröhlich, Selective Long-range Dispersion Forces between Large Sustems, *Phys. Lett.* 39A:153 (1972).

10. H. Fröhlich, Evidence for Bose-condensation-like Excitation of Coherent Modes in Biological Systems, *Phys. Lett.* 51A:21 (1975).

11. H. Fröhlich, The Extraordinary Dielectric Properties of Biological Materials and the Action of Enzymes, *Proc. Nat. Acad. Sci.* 72:4211 (1975).

12. B.L. Crawford, Jr. and H.L. Dinsmore, Vibrational Intensities.I. Theory of Diatomic Infrared Bands, *J. Chem. Phys.* 18:983, 1682 (1950).

13. R.C. Herman and K.E. Shuler, Intensities of Vibration-Rotation Bands, *J. Chem. Phys.* 21:373 , 22:954 (1953).

14. K.H. Illinger, Ed., " Biological Effects of Nonionizing Radiation ", ACS Symposium Series 157 (American Chemical Society, Washington, DC, 1981.

15. W. Heitler, " The Quantum Theory of Radiation ", 2nd. Ed., Oxford University Press, London, Chap. 3.

16. S.J, Webb, M.E. Stoneham and H. Fröhlich, Evidence for Nonthermal Excitation of Energy Levels in Active Biological Systems, Phys. Lett. 63A:407 (1977).

17. S.S. Penner, " Quantitative Molecular Spectroscopy and Gas Emissivities ", Addison-Wesley, Reading, MA (1959), Chap. 1.

18. A.C. Albrecht, On the Theory of Raman Intensities, J. Chem. Phys. 34:1476 (1961).

19. G. Placzek, in " Handbuch der Radiologie ", E. Marx, Ed., Akademische Verlagsgesellschaft, Leipzig, Vol. VII, Part II, p. 205.

20. W. Grundler, F. Keilmann and H. Fröhlich, Resonant Growth Rate Response of Yeast Cells Irradiated by Weak Microwaves, Phys. Lett. 62A:463 (1977).

21. W. Grundler and F. Keilmann, Nonthermal Effects of Millimeter Microwaves on Yeast Growth, Z. Naturforsch. 33c: 15 (1981).

22. W. Grundler and F. Keilmann, Sharp Resonances in Yeast Growth Prove Nonthermal Sensitivity to Microwaves, Phys. Rev. Lett. 51:1214 (1983).

23. H. Fröhlich and F. Kremer, " Coherent Excitations in Biological Systems ", Springer Verlag, Heidelberg (1983).

24. K.H. Illinger, Cellular Effects: Millimeter Waves and Raman Spectra, Proc. IEEE 68:114 (1980).

25. K.H. Illinger, Electromagnetic-Field Interaction with Biological Systems in the Microwave and Far-Infrared Region, in Ref. 14, p. 1 (1981).

26. F. Drissler, Discovery of Phase Transitions in Photosynthetic Systems, Phys. Lett. 77A:207 (1980).

27. K.H. Illinger, Spectroscopic Properties of *in vivo* Biological Systems. Boson Radiative Equilibrium with Steady-State Non-Equilibrium Molecular Systems, Bioelectromagnetics 3:9 (1982).

SPECTROSCOPIC STUDIES OF NERVES

B. Simic-Glavaski

Chemistry Dept. & Case Center for Electrochemical Sciences
Case Western Reserve University
Cleveland, Ohio 44106

ABSTRACT

Light-scattering techniques were used in studies of nerves, both at rest and at different levels of steady-state polarizations. The techniques employed quazi-elastic autocorrelation, Raman and infrared spectroscopies. Observations were made on the walking leg nerves of the lobster (*Homarus americanus*) and the connectives between the cerebropleural and buccal ganglia of a sea slug (*Pleurobranchaea californica*). Autocorrelation spectroscopy indicates that the scattering particles are time-independent. The properties of the autocorrelation functions are related to the degree of applied polarization and indicate a change in the scattering diameter as a function of the polarization. Raman spectroscopy indicates vibrational bands which closely resemble the Raman spectrum of ribonucleoprotein complexes. Infrared spectrum from a denatured nerve system is similar to the one obtained by Raman spectroscopy and indicates phospholipid structures. Influence of probe dyes on nerve systems were examined and changes were reversible.

INTRODUCTION

During the past years, conceptualization of the cell membrane has increasingly emphasized the membrane's dynamic fluid model in which metabolically selective molecules are dispersed or incorporated into a membrane matrix.[1] This conceptualization has been proven by a great body of experimental data, and, in turn, has provided a number of new techniques in studies of nervous systems. Current and voltage noise spectra[2,3,4] have yielded considerable information about the selective ion gates of neurons. However, traditional electrophysiological techniques have been based on invasive aspects of the microelectrodes. Light, as a molecular probe, has been used in studies of nerves[5,6] with an aim to provide physical and chemical information concerning molecular behavior in the neuroplasm and membrane.

Spectroscopic studies of nerves have used quasi-elastic light scattering (homodyne and heterodyne) and are of particular interest. By using these techniques, Shaw and Newby,[7] and then Piddington and Sattelle,[8] first obtained evidence of possible cytoplasmic motion. In studies of nerve axons, it was shown that light-scattering spectra are voltage-dependent and that the Hodgkin-Huxley formalism of photic noise is similar to that of voltage fluctuations.[9,10] However, there still remains a challenging task to provide an experimental technique for in situ and non-invasive studies which can reliably monitor molecular activity in a nerve membrane during its voltaic perturbation.

The other important question is associated with a non-destructive analytical methodology which can provide information about molecular specificity and their activity in complexes which may constitute ionic gates.

Motivated by such conceptual aspects, one can employ quasi-elastic and Raman spectroscopies in studies of nerve systems and provide some information about the action potential at a molecular level. Preliminary data of the quasi-elastic and Raman spectra obtained from the nerve bundles from the walking legs of the lobster and connectives between the cerebropleural and buccal ganglia of *Pleurobranchaea californica* are discussed in this report.

METHODS AND PROCEDURES

Nerve bundles from the walking legs of the lobster *Homarus americanus* and connectives between the cerebropleural and buccal ganglia of marine gastropod *Pleurobranchaea californica* were used in present experiments. The nerve bundles from the walking legs of the lobster contain large sensory and motor axons. The nerves can be readily dissected and usually remain physiologically active for many hours. The walking nerve neurons are surrounded by little connective tissue and are a standard neurophysiological preparation. In contrast to the walking leg nerves, cerebrobuccal connective is surrounded by an easily removable hemocoel and a very thin but tenacious connective tissue sheath that encapsulates the entire nervous system. This preparation can be neurophysiologically viable for several days after dissection.

Nerve preparations, dissected and isolated from muscle and connective tissue, were mounted in a physiological and optical lucite cell (Fig. 1). The cell was made with three section grooved to accept the nerve as it was gently laid across the compartments. The compartments were isolated by vaseline and then filled with artificial sea water. The central compartment, about 3 mm across, was easily accessible, and the nerve section in it was illuminated with the laser beam. The scattering volume was about 2 x 2 x 2 µm, and the scattered light was usually collected under 90°. Each compartment was supplied with a silver chloride electrode which was used for electrical polarization and recording of the electrical signal. Experiments were conducted at room temperature. In quasi-elastic light-scattering experiments, the source of light was a He-Ne SpectraPhysics laser, Model 124B, with a vertically polarized output of 25 mV at 632.8 nm.

In Raman experiments, an argon ion laser Coherent Radiation, Model 4G, was employed. The radiation was at 488.0 and 514.5 nm. The nerve samples could be illuminated with as much as 700 mW of the laser output for several hours and did not show any appreciable change in the conduction velocity or the shape of the action potentials evoked by electrical stimulation of one end of the nerves. However, much smaller power was used in the present experiments. In quasi-elastic measurements, one measures the time-dependent intensity fluctuation of the scattered light. The apparatus consisted of an ITT FW 130 photomultiplier whose output was fed into an amplifier-discriminator and then into a 64-channel Langley-Ford digital correlator. The generated autocorrelation function was displayed

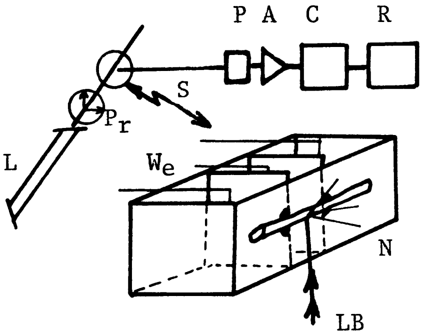


Fig. 1: Experimental arrangement for quasi-elastic light scattering from nerves: L, laser; P_r, polarization rotator; S, scattering chamber; N, nerves; LB, laser beam; W_e, working electrodes; P, photomultiplier; A, amplifier-discriminator; C and R, 64-channel correlator and recorder.

by an ASR33 Teletype and X-Y plotter. In quasi-elastic light-scattering experiments, one measures an autocorrelation function which, for a monodispersed system, is given by the expression:

$$R(\tau) = Ae^{-\Gamma\tau} \tag{1}$$

where A is a constant and Γ is a decay factor related to the molecular translational diffusion D and the wave vector $\vec{K}$ of the scattering system.

$$\Gamma = 2\ D\ K^2 \tag{2}$$

$$\vec{K} = (\frac{4\pi}{\lambda})\ n\ \sin\frac{\Theta}{2} \tag{3}$$

λ, n and Θ are the wavelength of the incident laser light, the refraction index of the medium and the scattering angle, respectively.[11]

To the first approximation, one can employ the three-dimensional Stokes-Einstein equation and relate molecular diffusion to the radius of scattering particles through the expression:

$$D = \frac{kT}{6\pi\eta r} \tag{4}$$

where k is the Boltzmann constant, T is the temperature and η is the viscosity of the medium in which scattering particles of the radius r randomly move.

In Raman scattering experiments, the nerve fibers were illuminated from the bottom and the scattered light was collected at 90°. A double monochromator Spex 1400 photomultiplier ITT FW 130 and a photon counting system were used. Averaging techniques in a multichannel mode were used with the aid of multichannel analyzer, Nuclear Data Model 600.

Results and Discussion

Quasi-elastic light-scattering data and the corresponding correlation function obtained from the nerve bundles from the lobster and *Pleurobranchaea* connectives are shown in Fig. 2a and 2b. Figure 2 (c1) illustrates also the corrleation function from a control experiment of 110. nm in diameter latex monodispersed spheres in distilled water. It is interesting to note that the correlation functions monotonically decay and can be best represented by single exponential fit as shown in Fig. 2 (a2, b2, c2). The experimental data obtained at constant potentials of the nerve fibers indicate scattering particles of uniform size. In addition, if one plots the decay shape Γ as a function of the scattering angle, then one

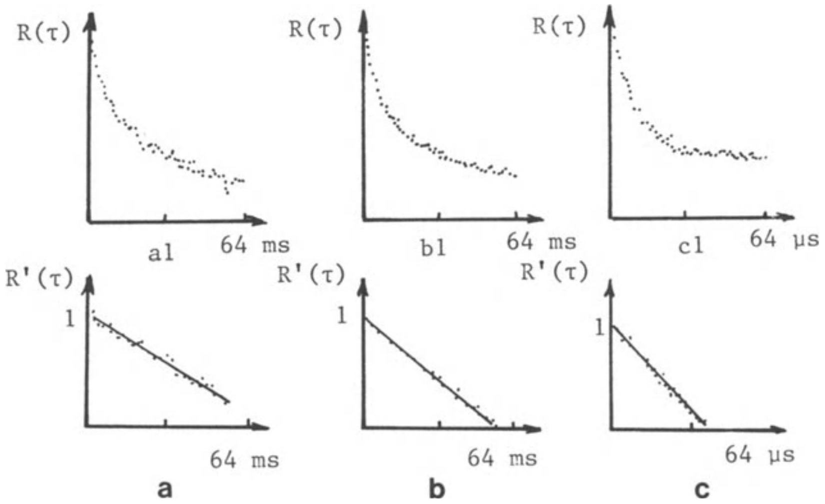


Fig. 2: Top Curves: Autocorrelation function R(τ) of vertically polarized light scattered from: lobster walking leg nerves, a1; cerebropleural-buccal ganglia connectives of *Pleurobranchaea*, b1; latex 1100 Å diameter spheres, c1. Bottom Curves: Normalized autocorrelation function $R(\tau)=\log[R(\tau)-R_b/R(o)-R_b]$, where R_b is the baseline scattering level. Solid line shows the best computer fit for single exponential data points. Time axis, 1 ms per channel in a and b, 50 μs in c. Autocorrelation functions were obtained at the nerves' resting potential.

obtains a linear relationship between Γ and the scattering angle as shown in Fig. 3. Thus, while differences in the correlation functions could arise from a variety of causes, such as association or disassociation of particles or even their conformational changes, the linearity illustrated in Fig. 3 suggests that they arise instead solely from the diffusion properties of the scattering particles, and, moreover, that the size and shape of these particles at constant membrane potentials are time-independent. It was found experimentally that correlation functions are greatly affected by applied electric fields on the nerve samples. The correlation functions were obtained at different electrical potentials both in hyperpolarizing and depolarizing directions. Constant voltages were applied between the central compartment and the two adjacent ones. The slopes of the correlation function are functions of the applied potential and diffusion constants are calculated for various potential values and shown in Fig. 4 for the *Pleurobranchaea* and lobster nerves.

If one uses appropriate values of the parameters in the Stokes-Einstein equation, it is now possible to calculate a radius of diffusing molecules.

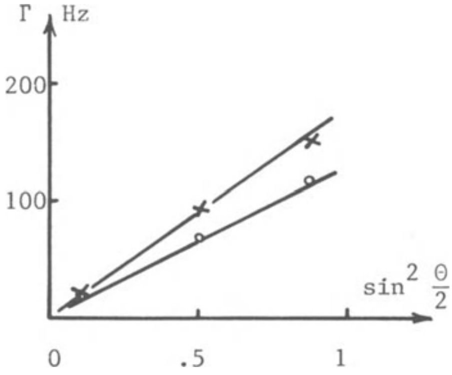


Fig. 3: Γ of correlation function versus scattering angle. (x), lobster; (∘), *Pleurobranchaea*.

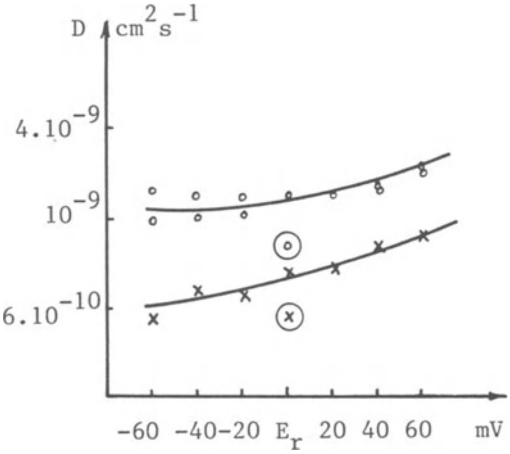


Fig. 4: Translational diffusion coefficient D as a function of depolarizing (positive) and hyperpolarizing (negative) potentials; E_r is resting level; (x) are data points for lobster nerves; (∘) are for *Pleurobranchaea*. Circled points were obtained after nerves were bathed in 5×10^{-5} M methyl green for 20 minutes. Recovery was to original value after a 12-hour wash.

Reported values of viscosity of 2 poise[12,13] for lobster nerve membranes and measured values of the diffusion constant indicate scattering particles of about 9 nm in the radius. Experimentally measured diffusion constants for lobster nerves and connectives are about 10^{-9} cm^2/s and are very well in agreement with values reported earlier for diffusing macromolecules in membrane systems[14].

Moreover, the change of the diffusion constants as a function of the applied potential show a change of molecular size of 6.7 nm/70 mV. This change of 6.7 nm may be closely related to the molecular association processes or their conformational changes most probably associated with the surrounding water molecules. Probe dye molecules have been frequently used as an optical readout in studies of the dynamic properties of neurons especially in recordings of the action potential. However, this experimental work indicates that considerable attention must be paid to the effects that such dyes can promote. Staining the nerve systems with dye molecules usually creates complexes which significantly perturb the original molecular system. In Fig. 4 are displayed changes of the diffusion constants for the lobster and *Pleurobranchaea* nerves when they were exposed to methyl green, $5 \cdot 10^{-5}$ M in the bathing medium. Interaction of dye molecules with the nerve molecules has a reversible effect and shows a change in diffusion constant equivalent to about 60 mV of applied polarization. Other dyes, such as methyl orange and red, have similar effects. Though dyes as photic probes can be used to detect action potentials, one should be aware of their interaction mechanisms which may significantly alter molecular conductivity properties.

While quasi-elastic scattering follows molecular hydrodynamic properties, one still needs information concerning specific molecular involvement in the action potential. To this end, Raman spectroscopy appears sufficiently sensitive to provide such specific answers, and, most importantly, can be applied in situ and for aqueous media where infrared spectroscopy seriously fails.

In Raman scattering experiments, laser excitation lines were at 488 and 514.5 nm. The scattering is accompanied with appreciable background of about 5000 counts. The spectra were recorded between 1100 and 1900 cm^{-1}. It was necessary to perform averaging, and 23 scans were used to obtain the Raman spectrum from lobster nerves, displayed in Fig. 5a. Eight Raman bands could be resolved in the observed spectral region.

An independent study of the nerve was also performed by using Fourier Transform infared spectroscopy. The nerve material was ground and deposited as a dehydrated film on silver bromide plate. Two hundred scans were preformed between 450 and 3800 cm^{-1} with the tesolution of 2 cm^{-1}, shown in Fig. 5b.

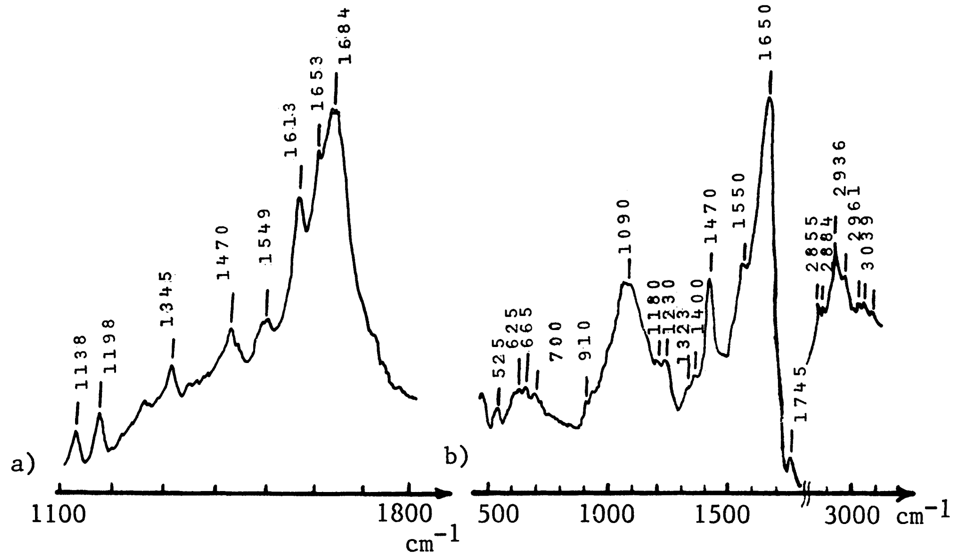


Fig. 5: a) Raman spectrum from the lobster nerve, 23 scans, resolution 5 cm^{-1} and 90 minutes recording time. Recording at the nerve resting potential. Laser excitation is 100 mW at 488. nm.
b) Infrared spectrum from casted nerve film, 200 scans, resolution 2 cm^{-1}.

Similarity between these two spectra is obvious and analysis of the appropriate bands is summarized in Table 1. A comparative analysis of spectra leads to an interesting similarity between the spectra obtained in this study and previously obtained spectra from aqueous solutions of RNA[15] and phospholipid-dipalmitoyl phosphatidycholine[16].

Müller and Rudin[17], and also Hyden,[18] suggested involvement of RNA complexes in the nerve system. According to Hyden,[18] neurons and the glia cells are rich in RNA and may amount to from 5 to 10% of their dry weight. It is quite conceivable that the large portion of the Raman signal is obtained from the glia cells which surround the nerves. Slight differences between the band positions, obtained in this study, and Raman bands obtained from the aqueous solution of RNA, may be accounted to the base stacking interactions.

The high frequency region between 2800-3100 cm^{-1} in the infrared spectrum is very well in agreement with those values obtained by Gaber, et al. from the solid phospholipid-dipalmitoyl phosphatidylcholine, indicating C-H stretches and a large content of phospholipids in the lobster nerve.

TABLE I

Raman (cm^{-1})	Infrared (cm^{-1})	RNA (cm^{-1})[15]	Phospholipid (cm^{-1})[16]
	This work		
	525 disulfide stretch		
	625 sec amide VI		
	665 amide V disordered polypeptide		
	700		
	910 C-H diform.		
1138	1090		
1198	1180	1185 A.U.G.C.	
	1230		
1345	1323	1340 U.	
	1400		
1470	1470	1484 U.	
1549	1550	1527 A.C.	
1613			
1653	1650	C-O stretch	
1684	1745 triglyceride		
	2855		2849 $\overset{H}{\underset{H}{C}}$
	2884		2880 $\overset{H}{\underset{H}{C}}$
	2935		2935 $\overset{H}{\underset{H}{C}}{-}H$
	2961		2961 $\overset{H}{\underset{H}{C}}{-}H$
	3039		3039 $N{-}\overset{H}{\underset{H}{C}}{-}H$

U, C, A and G indicate characteristics of the uracil, cytosine, adenanine and guanine bases, respectively.

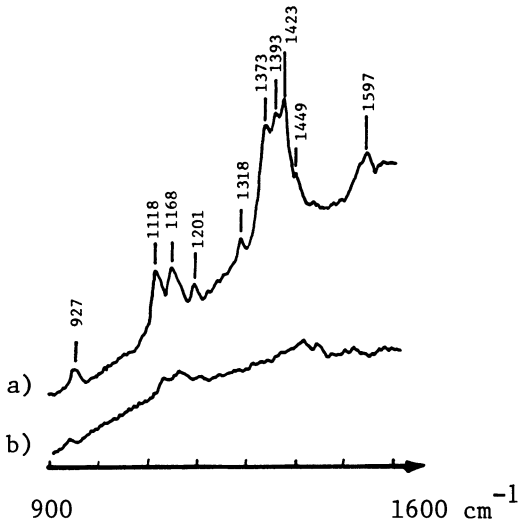


Fig. 6: a) Resonant Raman spectrum from $5 \cdot 10^{-5}$ M methyl orange in sea water
b) Raman spectrum from lobster nerve stained by $5 \cdot 10^{-5}$ M methyl orange for 15 minutes and then washed several times by clear sea water. Recording at the nerve resting potential. Laser excitation is 100 mV at 488. nm. Resolution: 2 cm^{-1}.

The translational diffusion data displayed in Fig. 4 have shown a significant change in the diffusion constant obtained from bathed nerves in methyl green solutions. This change in the diffusion constant suggests most probably an interaction of dye molecules with electrically active gate molecules in the nerve membrane. Similar effects were studied and observed by Raman spectroscopy. An aqueous solution of methyl orange shows resonant Raman effects when the solution was illuminated by the laser line at 488. nm.

The resonant Raman spectrum of $5 \cdot 10^{-5}$ M methyl orange in sea water solution was recorded and shown in Fig. 6a.

One scan was sufficient to provide an intense signal of about 65000 counts. The walking leg lobster nerve was stained for 15 minutes in the solution of sea water with $5 \cdot 10^{-5}$ M of methyl orange. Then the nerve was washed and placed in a fresh sea water bathing medium. Even after several washings, the nerve had a relatively strong yellow color, visible to the eye. The Raman spectrum was then recorded from the stained nerve, shown in Fig. 6b. Methyl orange bands at 1415 and 1423 cm^{-1} originate from an azo group symmetrical stretching mode N=N, and bands at 1118, 1168 and 1201 cm^{-1} originate from S=O and skeletal π electrons.[19] These bands change drastically in intensity and position in the dye-nerve spectrum and indicate loss of resonant effect in methyl orange and strong interaction with membrane molecules. Carey et al., in their study of methyl orange-albumin interaction, believe that dye molecules penetrate into protein molecules. However, diffusion data obtained in this work indicate an association interaction mechanism without penetration of the dye molecule.

In conclusion, the applied light-scattering techniques to *in situ* studies of nerve systems appear quite promising. One can gain some insight into the molecular properties of membranes that might be used as naturally exciting probes for photic monitoring of the steady-state and dynamic properties of neurons in order to eliminate as much as possible the invasive aspects of the more traditional electrophysiological techniques.

These initial studies encourage more extensive examination with neuron preparations that can be perfused and electrically clamped.

ACKNOWLEDGEMENTS

The author wishes to express his thanks to the NSF and to Professors G.J. Mpitsos and P.D. Junger for constant encouragement and fruitful discussions.

REFERENCES

1. S.J. Singer and G.L. Nicolson, Science 175:720 (1972).
2. H.M. Fishman, Proc. Nat. Acad. Sci., U.S.A. 70:876 (1973).
3. F. Conti, L.J. DeFelice and E. Wanke, J. Physio. 248:45 (1975).
4. R.J. Van den Berg, E. Siebenga and G. DeBruin, Nature, 265:177 (1977).
5. L.B. Cohen, R.D. Keynes and J. Hille, Nature 218:438 (1968).
6. I. Tasaki, A. Watanabe, R. Sadlin and L. Varnay, Proc. Nat. Acad. Sci., U.S.A. 61:883 (1968).
7. T.I. Shaw and B.J. Newby, Biochim. Biophys. Acta 225:411 (1972).
8. R.W. Piddington and D.B. Sattelle, Proc. Royal Soc. B. 190:415 (1975).
9. L.E. Moore, M. Tufts and M. Soroka, Biochim. Biophys. Acta 382:286 (1975).
10. B. Simic-Glavaski and L.E. Moore, Phys. Soc. Europ. Study Conf., Verbier, Switzerland (1974).
11. H.A. Cummins, in "Photon Correlation and Light-Beating Spectroscopy", H.A. Cummins and E.R. Pike, eds., pp. 285-330, Plenum Press, New York (1974).
12. M. Edidin, Ann. Rev. Biophys. Bioengin. 3:;79 (1974).
13. W.L. Hubel and H.M. McConnel, Proc. Nat. Acad. Sci., U.S.A. 63:16 (1969).

14. J. Schlessinger, D.E. Koppel, D. Axelrod, K. Jacobsen, W.W. Webb and E.L. Elson, Proc. Nat. Acad. Sci., U.S.A. 73:2409 (1976).
15. G.J. Thomas, Biochim. Biophys. Acta, 213:417 (1970).
16. B.P. Gaber and W.L. Peticolas, Biochim. Biophys. Acta, 465:260 (1977).
17. P. Müller and D. Rudin, Nature, 213:603 (1967); J. Theor. Biol. 18:222 (1968).
18. H. Hyden, in "The Neurosciences", G.C. Quarton, T. Melnechuk and F.V. Schmidtt, eds., pp. 248, New York, Rockefeller Univ. Press (1967).
19. P.R. Carey, H. Schneider and H.J. Bernstein, B.B.R.C., 47:588 (1972).

THE MODIFICATION OF DAVYDOV SOLITONS BY THE EXTRINSIC H-N-C=O GROUP

Scott P. Layne

Center for Nonlinear Studies
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

INTRODUCTION

The molecular mechanisms which underlie general anesthesia are not clearly understood. First, there are two chemically disparate classes of pharmacologic agents to consider (both intravenous and volatile) which induce general anesthesia. Second, investigators in the field are divided into two differing camps of thought: (1)those who believe that anesthetics work by altering normal membrane fluidity, and (2) those who believe that anesthetics work by perturbing normal protein function. Recently, there is growing evidence that the "perturbed protein" hypothesis holds greater promise over the "altered fluidity" hypothesis in explaining the molecular mechanisms of general anesthesia.[1] The simplest working idea is that general anesthetics act by binding directly to a particularly sensitive protein, which may or may not be located in a lipid membrane, and inhibiting its normal function.

Our proposed mechanism for anesthetic activity is a particular model of the "perturbed protein" hypothesis. We propose that intravenous agents are capable of changing localized structures within proteins and thereby altering normal protein dynamics. We will concentrate on a specific class of pharmacologic compounds which contain the H-N-C=O (or amide-I) moiety, rather than on volatile agents such as Halothane and Enflurane. Barbiturates, the most common intravenous agent in this class, will be the focus of attention. Our model also shows how several similar classes of agents called hydantoins, succinimides, glutethimides and urethanes are capable of perturbing normal protein function.

Questions which arise naturally from the "perturbed protein" hypothesis are: 1) what functions of proteins are disturbed by intravenous anesthetic agents, 2) which proteins are disturbed during general anesthesia, and 3) are certain proteins by virtue of their location in a cell more important than others?

(1) Our basic assumption is that intravenous anesthetics work by altering both the energy organizing and transporting functions within proteins. This idea, while around for a long time, has remained vague and unverified in the absence of a concrete model of how proteins perform these functions normally. This unsatisfying situation may have changed in 1973 when Davydov and Kislukha,[2] proposed a mechanism whereby energy could be transported efficiently along a one-dimensional molecular chain. In a paper entitled "Solitary Excitations in a One-Dimensional Molecular Chain," the authors considered how energy could be transferred via a coupling of intramolecular excitations and intermolecular displacments. In particular, the coupling of amide-I vibrational excitations to hydrogen bond displacements, and vice versa, leads to the production of a solitary wave in the one-dimensional molecular chain which does not disperse. Since 1973, Davydov[3] and his co-workers have published a number of papers on solitons in alpha-helical proteins. They address the important property of efficient energy transfer by proteins over distances that are large in terms of biological dimensions. The soliton model of protein dynamics deserves the careful attention of biologists. Clearly, it does not answer every question about protein behavior. Nevertheless, it is motivating new questions and experiments in protein dynamics. In this article, the soliton model will be used as a paradigm for understanding the molecular mechanism of general anesthesia. More specifically, we propose that anesthetics interfere with normal protein dynamics by disturbing soliton propagation in these molecules.

(2) Assuming that the dynamical behavior of proteins may be perturbed by an intravenous anesthetic we concentrate, for the sake of simplicity, on the alpha-helical protein. This emphasis on the alpha-helix is justifiable, because the helical configuration is the most common of protein structures. Alpha-helical proteins are found in both the membrane and cytoskeletal fractions of cells. It has been estimated, on the basis of circular dichroism studies, that up to one third of membrane proteins are in the helical configuration, with the remainder of proteins in a random coil form.[4] Helical membrane proteins usually consist of hydrophobic amino acids and have a length in the range of 30-70 Å (5-10 turns). The helix is also a common structural motif in the cytoskeleton, where there are at least three important classes of contractile proteins that are highly alpha-helical: spectrin, tropomyosin and myosin. In addition to the contractile system, the extensive network of intermediate filaments in the cytoskeleton also demonstrates predominant helical character.[5] Cytoskeletal proteins tend to be much

longer than membrane proteins (300-1600 Å) and they also tend to form highly stable coiled-coil structures which are capable of polymerizing in a head-to-tail configuration.[6,7] This end to end configuration creates exceptionally long alpha-helical networks within the cytosol.

We will discuss how hypnotic agents which contain the amide-I moiety, such as barbiturates, are capable of perturbing solitons in alpha-helical proteins. This could occur via insertion of the anesthetic agent into the spines of the alpha-helix. This, in turn, modifies the intramolecular excitation and intermolecular vibrational properties of the protein. This perturbation of energy transduction through the helix involves merely a modification of structure in one portion of a helix, rather than a global conformational change. We will show by numerical results that if the propagation of a soliton is disrupted by an anesthetic molecule

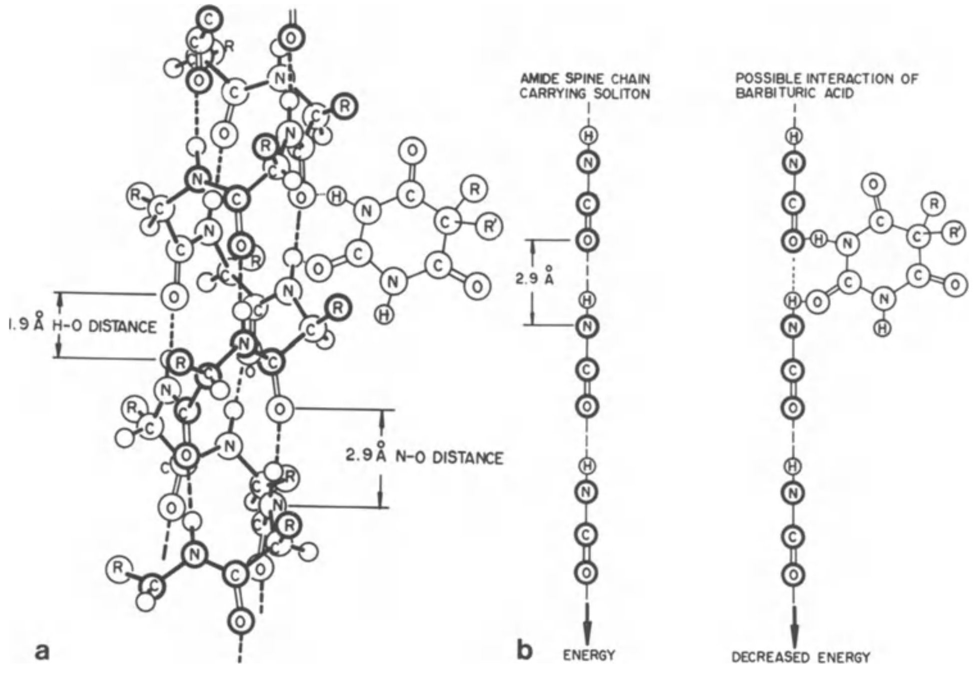

1.9 Å H-O DISTANCE
2.9 Å N-O DISTANCE
a
AMIDE SPINE CHAIN CARRYING SOLITON
2.9 Å
b
ENERGY
POSSIBLE INTERACTION OF BARBITURIC ACID
DECREASED ENERGY

Fig. 1. (a) One of barbiturate's possible interactions with an alpha-helical protein via its H-N-C=O moities. The spiral configuration of the protein is stabilized by its weak hydrogen bonds and the binding of a barbiturate changes the localized structure within the helix. (b) Isolated view of barbiturate's two-point interaction with the amide-I spine system that supports solitons in the alpha-helix. Such an interaction alters the critical one-dimensional configuration and hydrogen bond relations in the system.

at one bond then the movement of the soliton throughout the rest of the molecule is altered.

(3) We believe that two biological membranes are the principal sites for anesthetic activity: 1) the cellular (neuronal) membrane, and 2) the mitochondrial (inner) membrane. Although there are other sizable membranes within a cell, such as nuclear membrane, endoplasmic reticulum and golgi apparatus, these bilayers are not thought to be important to the mechanism of general anesthesia. Only membrane proteins that are related directly to intercellular communication and energy production will be considered as the primary targets for anesthetic agents. This assumption does not preclude the possibility that other membrane proteins are inhibited by hydrogen bonding anesthetic agents but these effects will be considered as secondary in the production of unconsciousness.

In addition, it should be mentioned that cytoskeletal proteins may be inhibited by hydrogen bonding anesthetic agents. The importance of this effect is rather controversial, since little is known about the energetics of this complex and heterogeneous network of proteins. Clegg[8] has speculated that the extensive cytoskeletal framework plays a greater role than simple structural support and cellular motility. He proposes that the cytoskeleton coordinates the energetics of the intracellular space. The cytoskeleton may provide a matrix on which the so-called soluble enzymes attach and coordinate their energetic activities. Cytoskeletal proteins are known to be associated closely with both the cellular and mitochondrial membranes. Membranes may "talk" to the cytoskeleton and vice versa. Therefore, complete models of general anesthesia should consider the interrelations between these anatomically distinct regions within neurons.

Green[9] has categorized biological energy into two broad and encompassing categories: 1) valence vibrational energy, and 2) gradient or ionic potential energy. As already suggested, hydrogen bonding anesthetic agents are capable of interfering with the propagation of the first category of biological energy. This action may occur at three possible sites. At the neuronal membrane, such an action will result in a diminution of communication between neurons. Decreased axonal conduction, decreased neurotransmitter action and increased release, and decreased synaptic signaling all point to the fact that membranes are less talkative and sensitive to their neighbor's messages.[10] This inhibitory action may involve both the protein and the substantial glycoprotein fractions on the neuronal membrane. At the mitochondrial membrane, anesthetic action will result in a diminution of ATP synthesis. It has been shown experimentally that barbiturates inhibit mitochondrial activity by acting between the iron-sulfur centers of complex I and coenzyme Q.[11] In other words, the anesthetic molecule stops oxidative phosphorylation at the very beginning of the ATP synthesis process. This location of activity is consistent with the hypothesis that barbiturates inhibit energy transport by proteins, since complex I contains the first organization of proteins which are reputed to participate in energy coupled processes and since a shutdown at complex I will halt the majority of activities of the tripartite repeating unit. At the cytoskeleton, such an action may result in a decreased coordination between the intricate intertwinings of the long and filamentous proteins. Apparently, the shape of the cytoskeleton is dependent on the various micro-environmental concentrations of ATP and Ca^{2+} within the neuron.[5] Barbiturates may interfere with these subtle processes and such a loss of coupling from within could alter the cytoskeleton's response to membrane commands. Gradient energy, or the separation of charge across a lipid bilayer, can also be perturbed by hydrogen bonding anesthetic agents. This is understandable because the second category of biological energy is basically a manifestation of valence vibrational energy. The movement of charge across a membrane must first be initiated by an energetic change in the structure of proteins.[12] If a soliton participates in this charge separating operation by proteins then gradient energy can be affected.

In the brain consciousness can be likened to the "frosting on the metabolic cake." In other words, a relatively small cutback in total metabolic output of the CNS (about 10-15%) will result in unconsciousness. Therefore, it is necessary that barbiturates inhibit only a select percentage of total protein activities. This will result in a relative loss of communication between neurons which is manifest as a loss of consciousness. In the anesthetic state, the neuronal membranes are quiescent and demanding a smaller energy supply. This is balanced by a decreased synthesis of ATP by mitochondria so that the relative turnover at both membrane sites is diminished. This helps to explain the fact that during barbiturate anesthesia the brain settles down to a metabolic valley where demand for an energy supply meets production of the energy source.[13]

SOLITONS: A POSSIBLE MODEL FOR ANESTHETIC ACTIVITY

In the soliton model, the alpha-helical protein is approximated as three amide-I spines, in other words, three separate but parallel chains of repeating H-N-C=O groups. Vibrational excitations within the amide-I groups are coupled nonlinearly to molecular displacements along the chain and this nonlinear coupling can produce a soliton. In order for a soliton to travel through an alpha-helix, several basic requirements must be satisfied: 1. The amide-I spines must be sufficiently parallel to propagate the longitudinal sound and bond energy of a soliton. 2. The nonlinear hydrogen bond coupling between adjacent amide-I groups must be above threshold for cooperative interactions to take hold. 3. There must be sufficient vibrational energy for nonlinear coupling to take hold. Since barbiturates contain four H-N-C=O groups in a rigid planar structure, it is straightforward to see that this anesthetic molecule contains four amide-I groups which can hydrogen bond to an alpha-helix. Therefore, an interaction of a barbiturate with an alpha-helix might alter either the propagation or the formation of solitons.

The hydrogen bonding of a barbiturate to an alpha-helix will distort and break the intramolecular hydrogen bonds that are responsible for creating and maintaining the spiral structure of the protein. This hydrogen bonding may occur via a simple insertion as shown in Fig. 1, or it may involve a more complex geometry between anesthetic and protein molecules. For example, many helical membrane proteins are known to be combinations of both parallel and antiparallel alpha-helices. Cytochrome proteins are a good example of this configurational motif. Commonly four alpha-helices, which are ~ 40Å long and separated by ~ 5Å, line up together to create a larger up-and-down helix bundle.[14] It is likely, for a more complex anesthetic-protein interaction, that a barbiturate will occupy a variety of inter- and intramolecular positions within the cytochrome protein. These various positions may involve the breaking of hydrogen bonds within the H-N-C=0 chains of the cytochrome or they may involve interactions within the side chains of the helices that help to stabilize the overall up-and-down structure of the protein. Fortunately, in spite of the enormous variety of interactions between anesthetic and protein molecules, the disturbance of energy migration may still be investigated by looking at increasingly complex models of soliton propagation.

We gain some initial insight by considering the continuum approximation for soliton propagation on a single chain, namely, the nonlinear Schrödinger equation (NLS):

$$i\hbar A_t + A_{xx} + G_o|A|^2A = 0 \ , \tag{1}$$

where A is the amplitude of the wave and the nonlinear coupling constant $G_o = 4\chi^2/J\omega$. χ is the exciton-phonon interaction parameter, J is the longitudinal dipole-dipole coupling energy and ω is the longitudinal spring constant for the hydrogen bond. This continuum approximation is justified because we know from numerical calculations that a typical soliton spans eight helical turns and travels at 3/8 the sound of speed.[15,16] In others the amplitude varies continuously from one unit cell to another.

How would this equation be modified in the presence of an anesthetic agent? We believe that an anesthetic molecule creates a narrow region of distortion in the protein that does not denature the protein but does alter the value of nonlinear coupling in that region. Similar effects might be produced by interactions with side chains and neighboring residues which have so far been neglected in the model. These latter interactions are likely to produce changes in G_o over both wide and narrow regions. Suppose we allow G_o to vary as a function of x, such that $G(x) = G_o + \varepsilon G_1(x)$ and $\varepsilon \neq 0$. Calculations show that smooth variations in the value of G(x) allow a soliton to adjust its height and width to the changing nonlinearity and propagate with mild disturbance. However, relatively abrupt variations do not allow enough time for a soliton to adjust itself. Instead the shape and speed of a soliton are altered and it looses an increased amount of energy through radiation. The ratio of $G(x)/G_o$ and the width of G_x, when $\varepsilon \neq 0$, in relation to the width of a soliton are important considerations. The more this variation approaches a step function, the more soliton propagation is perturbed.

These qualitative insights were verified by our numerical studies of a more complicated model, the modified Davydov equations of Scott.[16] These equations have the form:

$$
\begin{aligned}
i\hbar \frac{da_{n\alpha}}{dt} &= [E_o + \chi_1(\beta_{n+1,\alpha} - \beta_{n-1,\alpha})]a_{n\alpha} - J(a_{n+1,\alpha} + a_{n-1,\alpha}) \\
&+ \chi_2[a_{n+1,\alpha}(\beta_{n+1,\alpha} - \beta_{n,\alpha}) + a_{n-1,\alpha}(\beta_{n,\alpha} - \beta_{n-1,\alpha})] \\
&+ LF_L + NF_N + PF_P + QF_Q + RF_R + SF_S + TF_T + UF_U + VF_V \\
&+ XF_X + ZF_Z \; . \qquad (2)
\end{aligned}
$$

$$
\begin{aligned}
M \frac{d^2\beta_{n\alpha}}{dt^2} - \omega(\beta_{n+1,\alpha} - 2\beta_{n,\alpha} + \beta_{n-1,\alpha}) &= \chi_1(|a_{n+1,\alpha}|^2 - |a_{n-1,\alpha}|^2) \\
&+ \chi_2[a^*_{n\alpha}(a_{n+1,\alpha} - a_{n-1,\alpha}) + (a^*_{n+1,\alpha} - a^*_{n-1,\alpha})a_{n\alpha}] \; . \qquad (3)
\end{aligned}
$$

The index n specifies a particular unit cell (H-N-C=0 group) counting longitudinally along the helix, while α specifies each of the three spines and takes the value 1, 2, or 3. The other terms have the meaning: $a_{n\alpha}$ is the probability amplitude for vibrational excitation of the amide-I group located at spines n,α; E_o is the quantum energy of an amide-I vibration; $\beta_{n\alpha}$ is the longitudinal displacement of the residue located at n, α; M is one-third the molecular weight of a unit cell; ω is the linear spring constant for a hydrogen bond; and J is the dipole-dipole coupling energy between a particular amide-I group and those ahead and behind along the same spine. The LF_L ... ZF_Z terms correspond to additional dipole-dipole coupling terms in order to account for helical structure. Finally, χ_1 orders the nonlinear coupling between exciton and phonon interactions (a more complete explanation follows) and χ_2 orders the change in the J term caused by stretching in the helix between n+1 and n.

Figure 2 shows the results of integrating these equations numerically for an unperturbed helix 900 Å long, consisting of three coupled chains of 200 H-N-C=0 groups. A nonlinear coupling constant $\chi_1 = 0.40 \times 10^{-10}$ newtons was assigned to every hydrogen bond in the molecule. This coupling constant is just above threshold for soliton formation, when two quanta of amide-I energy are used to initiate the pulse. The soliton was launched by placing a single quanta on spines $\alpha = 1$ and 2 at the first amide-I bond n=1. At computer time T=0, all other H-N-C=0 groups were at rest. This procedure ensured that in later calculations the soliton was well formed before it entered the region of perturbation. For all spines in the helix $J = 7.8\ \mathrm{cm}^{-1}$ (or 0.145 computer units) and $w = 19.5$ newtons/meter.

How will Eqs. (2) and (3) change in the presence of an anesthetic molecule? The simplest alteration is to change the values of J, w and χ_1.

J is the resonance interaction energy between adjacent amide-I dipoles and is approximated by the standard formula for two parallel dipoles

$$J \doteq \frac{\vec{d}^{\,2}}{4\pi\varepsilon_o R^3}\,(3\cos^2\theta - 1) \tag{4}$$

where $\vec{d}$ is the dipole moment, R is the distance between dipoles, ε_o is the dielectric coefficient and θ is the angle between dipoles. Since an anesthetic molecule may break the hydrogen bonds in an alpha-helix and cause it to unwind, the distance R between neighboring dipoles will increase and consequently J will decrease. The dipole-dipole coupling terms between neighbors ($NF_N \ldots ZF_Z$) will also be shifted by the distortion in helical shape but the major effects should occur near the point of anesthetic binding. Furthermore, the four amide-I moites in the barbiturate ring can dipole couple to the alpha-helix. This coupling will depend on the relative angles between the anesthetic and protein H-N-C=0 groups and in most instances this coupling will act to extract energy from a soliton. In the calculations below the only dipole constant that was changed was J. It was decreased from $7.8\ \mathrm{cm}^{-1}$ to $4.85\ \mathrm{cm}^{-1}$, which corresponds roughly to an increased distance R from 4.5 Å to 5.3 Å. From consistent force field calculations on the hydrogen bond,[17] an increase in hydrogen bond length of $\Delta R = 0.8$ Å corresponds approximately to a drop in hydrogen bond energy of 55%. This empirical decrease in J was chosen as a hopefully conservative estimate of an anesthetic interaction with a protein, since the hydrogen bond remains intact at this level of distorsion. More complete estimates should include changes

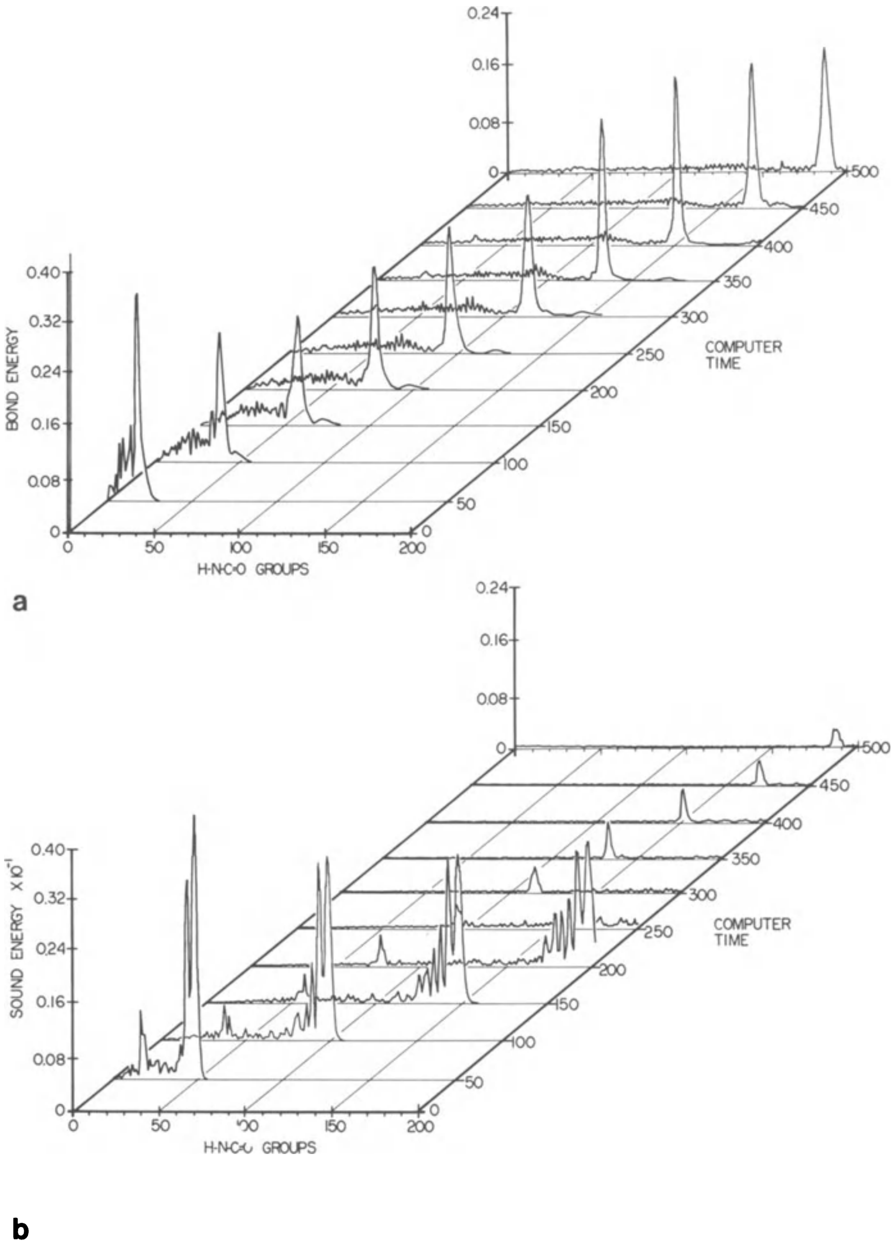


Fig. 2. Unperturbed soliton. When a soliton is initiated at one end of a helix, a typical amount of energy is not captured by the pulse. This energy moves at the sound speed and exits the helix by T = 225. The soliton consists primarily of bond energy (C=0 stretch) and the total energy that is carried as sound energy is ~ 0.04%. The energy scales correspond to the summed energy on all three spines. a) Is bond energy and b) is sound energy.

in all dipole-dipole coupling terms and energy extraction by amide-I groups in the anesthetic.

The spring constant ω for the hydrogen bond is related to the bond energy which, in turn, depends on the length and angle of the bond.[17] Longer and less energetic hydrogen bonds in proteins demonstrate lower spring constants which are associated with higher N-H and C=0 stretching forces. This means that upon hydrogen bonding -N-H → 0=C-, the energy of N-H and C=0 stretching declines.[18] The spring constant ω can also be considerably affected by changes in the N-H$\angle$0=C angle, since the strength of the hydrogen bond is rather sensitive to deviations beyond 20-30°. Therefore, an anesthetic interaction which unwinds and lengthens a hydrogen bond will diminish the spring constant for that bond. The extent to which this occurs is difficult to estimate accurately by simple methods. However, as a first approximation, our numerical study assumed a simple linear relationship between hydrogen bond energy and the spring constant ω. Hence, at $\Delta R = 0.8$ Å the restoring force $\omega = (19.5)(1.0-0.55) = 8.8$ newtons/ meter. This is a moderate estimate for a change in ω, since a completely broken hydrogen bond will have no intramolecular spring constant.

The exciton-phonon coupling constant "χ_1" can also be changed by the association of an anesthetic molecule with a helix. It is well known that the C=0 stretching mode occurs above 1700 cm^{-1} for a free carbonyl and that it is red shifted by hydrogen bonding.[19] Careri[20] points out that this red shift to ~ 1665 cm^{-1} is strictly proportional to the chemical shift of the proton as measured by NMR and this shift can be used as a good measure of hydrogen bonding. Thus χ_1 is the shift in energy of an amide-I vibrational quantum per unit change in hydrogen bond length

$$\chi_1 = \frac{dE_o}{dR} \tag{4}$$

where E_o is the energy of an amide-I vibration and R is length of the associated hydrogen bond. Scott[21] points out that $E_o \propto W^{\frac{1}{2}}$, where W is the spring constant of the C=0 vibration, which allows (4) to be written as

$$\chi_1 = \frac{E_o}{2W}\frac{dW}{dR} \tag{5}$$

When an anesthetic molecule binds to a helix and weakens (or breaks) its intramolecular hydrogen bonds the influence of a weaker hydrogen bond will cause dW/dR to decrease. This will affect a decline in the value of χ_1. Kuprievich and Kudritskaya[22] calculated a value of $\chi_1 = 3\text{-}5 \times 10^{-11}$ newtons and for the purposes of the present numerical study χ_1 is taken at this lower limit to approximate the coupling constant for a weaker hydrogen bond. According to the findings of

Scott[16], this value of χ_1 is just below threshold for soliton formation when two quanta of amide-I energy initiate the pulse.

We altered the numerical code described above to simulate two types of perturbations: (1) a smooth wide distortion (twice the soliton width) as might be produced by a side chain or neighboring residue, and (2) a narrow abrupt distortion (half the soliton width) as might be produced by a barbiturate. The smooth perturbation was centered at unit cell n = 105 of our 200 unit cell helix and was defined to smoothly decrease the values of J, ω and χ_1 from their unperturbed values. Specifically, $J_{n,\alpha}$ was given by:

$$J_{n,\alpha} = J_{max} - J_{cst} \text{ Sech } (n - 105)/K$$

where J_{max} is the unperturbed value of J, and $J_{cst} < J_{max}$. K sets the width of the perturbation seen at the top of Fig. 3. ω and χ_1 were defined similarly. The abrupt perturbation was introduced by a step function in the values of J, ω and χ_1 over the unit cells n = 100 - 103. In both cases the minimum values of the coupling constants were $\chi_1 = 0.3 \times 10^{-10}$ newtons, $J = 4.85$ cm^{-1} (or 0.09 computer units), ω = 8.8 newtons/meter, and as suggested above are reasonable estimates of a physically relevant perturbation.

The results of introducing these two perturbations are shown in Figs. 3 and 4. The value of $\chi_1^2/J\omega$ is plotted as a function of n in the upper corners of these figures. Notice that the measure of nonlinearity doubles at n = 105 for the smooth perturbation and at n = 100-103, for the abrupt perturbation. In both figures computer time $T = (M/\omega)^{\frac{1}{2}} = 0.99 \times 10^{-13}$ seconds and sound speed = 4.5 Å $(\omega/M)^{\frac{1}{2}}$ = 4.56×10^{13} Å/second. The time it takes a soliton to pass through one unit cell is 0.372 computer time units (T). This implies a soliton velocity of 1.69×10^{13} Å/second and a ratio of soliton to sound velocity of 0.37.

THE H-N-C=O MOIETY: A COMMON DENOMINATOR OF INHIBITORY ACTIVITY

Barbiturates demonstrate a broad range of lipid solubility. For example, thiopental has one of the highest oil-water partition coefficients at 4.7, whereas barbital has one of the lowest at 0.21. Factors that elevate the partition coefficient include both increased molecular weight (by C5 substitution) and decreased ionization at physiologic pH. It is well known that increased lipid solubility is associated with a shortened onset of action, a shortened duration of action, an enhanced hypnotic potency and an enhanced interaction with hydrophobic regions of protein.[23] This last property of increased hydrophobic interactions with proteins is especially important to the hypothesis of membrane activity. Factors which increase lipid solubility will also increase anesthetic-protein interactions at the neuronal and mitochondrial lipid membranes.

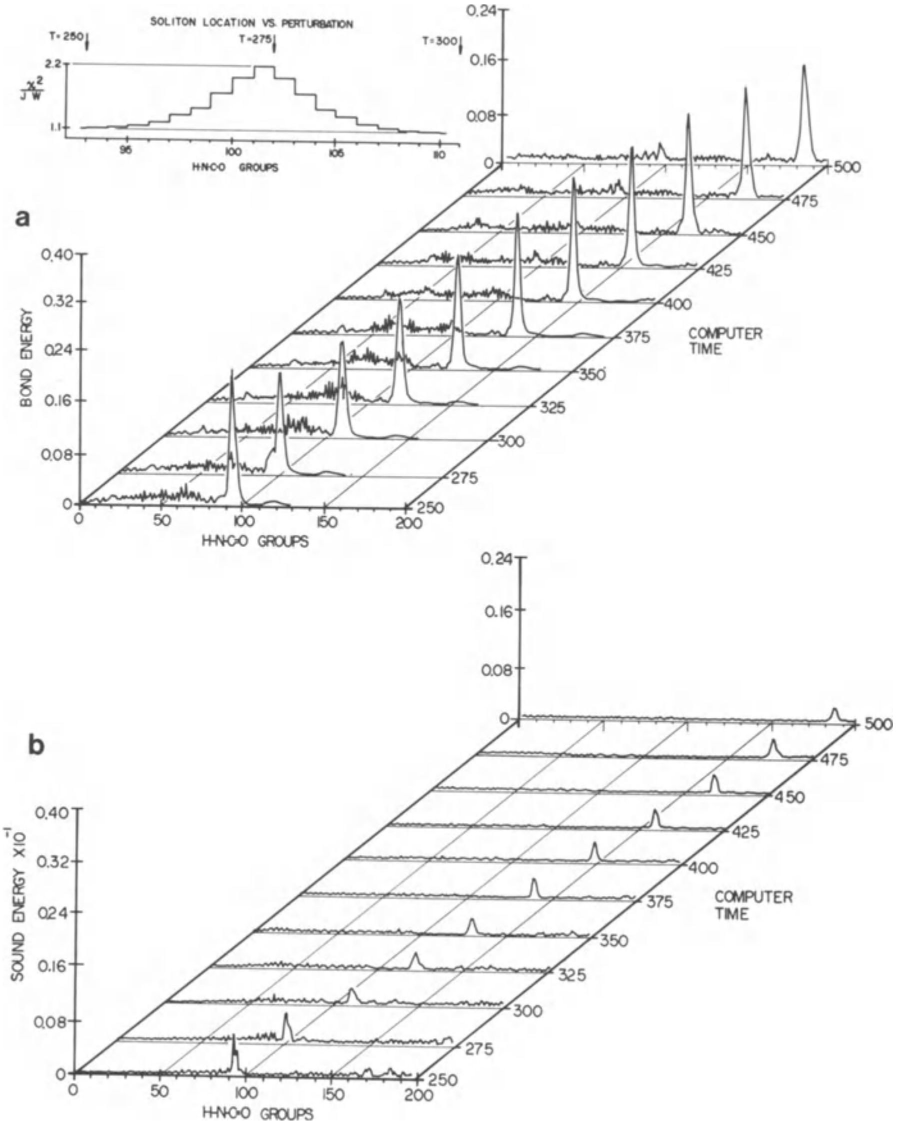


Fig. 3. Gradual disturbance. The smoothly changing value of $\chi_1^2/J\omega$ is shown by the upper insert. At T = 500 the height of the soliton is ~ 80% of the unperturbed pulse and the width is ~ 10% broader. Practically no energy is radiated by the soliton at the point of maximum perturbation T = 275. This ramp-like disturbance allows the soliton to adjust itself as it moves through the changing values of $\chi_1^2/J\omega$. Up to T = 250 soliton propagation is identical to Fig. 2. a) is bond energy and b) is sound energy.

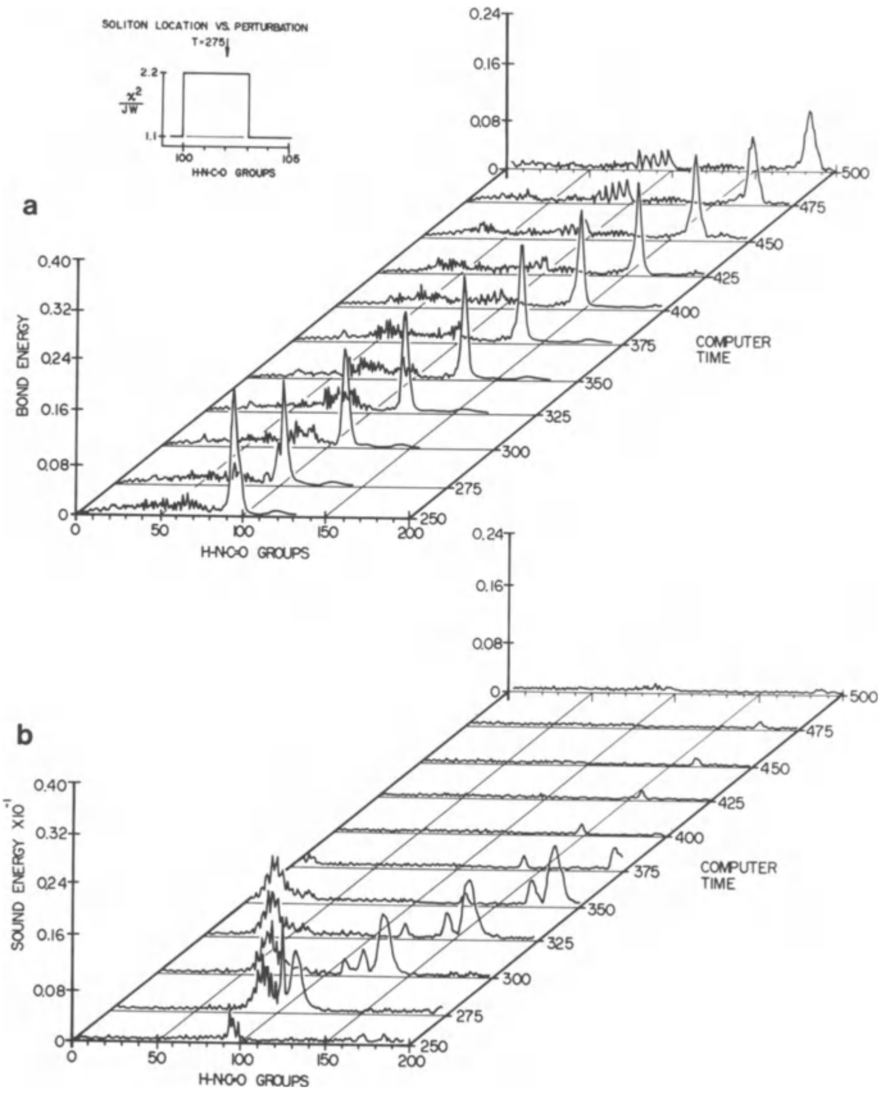


Fig. 4. Abrupt disturbance. The wall-like change in $\chi_1^2/J\omega$ is shown by the upper insert. At T = 500 the height of the soliton is ~ 50% of the unperturbed pulse and the width is ~ 20% broader. Also a tail has been formed which trails energy behind the soliton. At T = 275, increased sound energy is radiated by the soliton as it cuts through the region of perturbation. This abrupt disturbance does not allow the soliton to adjust itself as it moves through the changing value of $\chi_1^2/J\omega$. Up to T = 250 propagation is identical to Fig. 2. a) is bond energy and b) is sound energy.

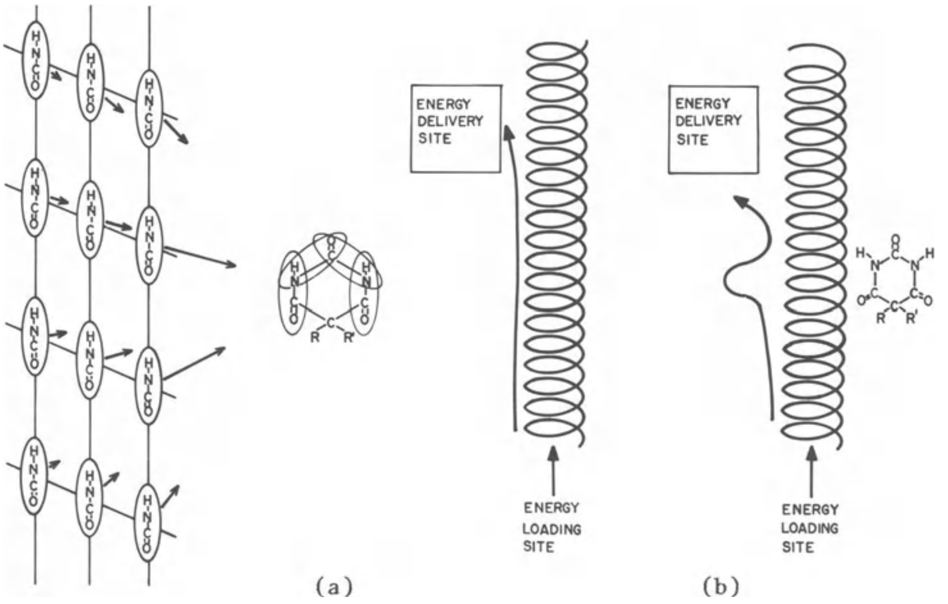


Fig. 5. (a) The interaction of barbiturate's four H-N-C=O dipoles with the dipoles of an alpha-helix. The protein is represented by a two-dimensional projection and the arrows indicate the relative energy of resonant interaction which is proportional to $1/R^3$ and dependent upon the angle between dipoles. (b) The propagation of a soliton can be disturbed by a barbiturate molecule. Thus when a soliton is launched, it does not make it to its energy delivery site.

To induce general anesthesia, the concentration of barbiturate in whole brain tissue must be on the order of 200 micromolar.[24] Since most anesthetically useful barbiturates demonstrate a lipid to water partition coefficient of ~ 4:1 and since whole brain tissue has more cytosol than lipid, it is reasonable to calculate a membrane concentration of $200 \times 4 = 800$ micromolar during general anesthesia. This concentration is equivalent to a volume distribution of $\sim 2 \times 10^6$ Å^3/barbiturate. Taking cytochrome C' as an average membrane protein,[14] which as mentioned earlier consists of four 40 Å helices separated by ~ 5 Å, implies a protein which encompasses a volume of about 20 Å x 20 Å x 40 Å = 1.6×10^4 Å^3/molecule. Therefore, during general anesthesia approximately $1.6 \times 10^4 / 2 \times 10^6 \times 100 = 0.8\%$ of typical membrane proteins are associated with an anesthetic molecule. This rough calculation assumes a uniform volume distribution and no increased affinity of hydrogen bonding anesthetics for proteins. It suggests a cooperative association among membrane proteins, since a possible inhibition of ~ 1% of membrane activities results in a 10-15% decrement in CNS oxygen consumption.

In addition to barbiturates, there are a number of classes of anesthetic, antiepileptic and hypnotic agents that have a similar H-N-C=O structure. The number of amide-I moities in these agents appears to be related both to potency and to duration of action. Hydantoins contain three, succinimides and glutethimides contain two, and urethanes contain one functional amide-I group (Fig. 6). Hydantoins, succinimides and trimethadione are commonly used as antiepileptic agents. Glutethimides are used as longer acting and urethanes are used as shorter acting hypnotic agents. N-methylated trimethadione is inactive until the hepatic microsomol enzymes demethylates the parent compound to the active N-H structure. Such structure-activity relationships support the proposals: 1) the entire H-N-C=O moiety is required for activity, and 2) the H-N-C=O moiety is involved in two-point hydrogen bonding. Both Sandorfy[25] and Rich[26] have found that barbiturates break the hydrogen bonds in model systems of purines and pyrimidines. This breaking of hydrogen bonds involves the ability of barbiturates (and related compounds) to form a hydrogen bonded cyclic dimer with the purine base. In conjunction with this, molecular conformational energy calculations have confirmed that the one-point hydrogen bonding abilities of barbiturates, hydantoins and succinimides is unrelated to either their hypnotic or antiepileptic activities. Further, these studies have suggested that hypnotic and antiepileptic agents require the participation of more than one carbonyl or amide hydrogen for the production of pharmacologic activity.[27,28] This suggestion of two-point H-N-C=O participation supports the present model of anesthetic activity.

In summary, there has not been a satisfactory model to account for barbiturate's diverse activities. Theories have implicated barbiturate's hydrogen bonding ability as a mode of action but they have failed to provide a specific context for activity. The soliton model offers a unifying mechanism of action for a number of diverse agents, such as, barbiturates, hydantoins, glutethimides, succinimides and urethanes. It accounts for the observation that barbiturates display no stereospecific activity at anesthetic concentrations since the one-dimensional spine system of an alpha-helix cannot be stereospecific. It accounts for the observation that N-methylated derivatives of barbituric acid are "ultrashort" hypnotics, since N-methylation will "choke" the total number of H-N-C=O configurations in the anesthetic ring. Finally, it relates the number of H-N-C=O groups in the molecule to hypnotic potency.

OPEN QUESTIONS

1. Our numerical studies of a relatively long alpha-helix do not necessarily apply to short helical proteins found in membranes. Therefore, similar numerical studies should be carried out to investigate these membrane proteins.

2. It is possible that a bartiburate might trap the soliton rather than extracting energy from it. This possibility should be investigated.

3. Rich[26] points out that barbiturates are capable of binding to adenine derivatives. Therefore, the amide-I groups in a barbiturate might absorb ATP hydrolysis energy before it can be delivered to proteins. This possibility should be considered in conjunction with the present model.

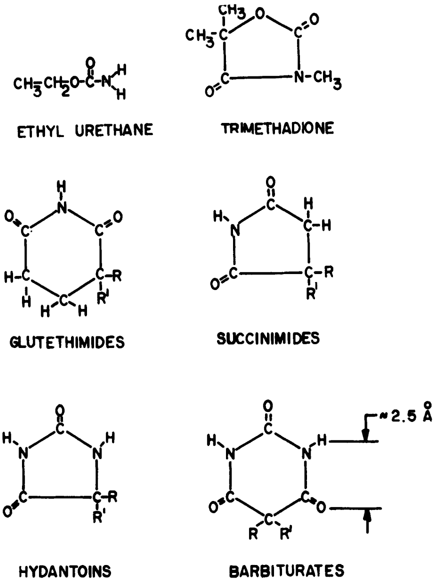


Fig. 6. Anesthetic and antiepileptic agents that are capable of disturbing solitons. The presence of an alkyl or aryl group at R and R' confers increasing lipid solubility and, generally, increased lipid solubility promotes an increased drug potency.

CONCLUSION

General anesthesia may be related to protein-anesthetic interactions which alters normal protein behavior. We believe that dynamical models of protein behavior are necessary for understanding such interactions. The soliton model may be a promising candidate for such an approach. It accounts for many of the pharmacologic properties of H-N-C=O containing compounds.

ACKNOWLEDGMENT

It is a pleasure to thank Peter Lomdahl for help with the numerical code. Thanks are also due to Ross Adey, Bob Birge, Dave McLaughlin and Al Scott for many stimulating discussions.

NUMERICAL METHODS

Details of the numerical computations are nearly identical to those given in references 15 and 16 except for the following points: The calculations were integrated using a leapfrog predictor-corrector method. For computations where there were changes in χ, J and ω values a time step was chosen to minimize the relative error per unit time interval. The accuracy of the time integration was checked by rerunning the calculations with a smaller time step and rerunning the calculations with a third order predictor second order corrector method with an equivalent time step. No significant differences were found between the smaller time steps and the two methods of integration.

REFERENCES

1. N. P. Franks and W. R. Lieb, Molecular Mechanics of General Anesthesia, Nature 300:487 (1982).
2. A. S. Davydov and N. I. Kislukha, Solitary Excitons in One-Dimensional Molecular Chains, Phys. Stat. Sol. 59:465 (1973).
3. A. S. Davydov, Biology and Quantum Mechanics, Pergamon, Oxford (1982).
4. J. Lenard and S. J. Singer, Protein Conformations in Cell Membrane Preparations as Studied by Optical Rotatory Dispersion and Circular Dichroism, Proc. Natl. Acad. Sci. 56:1828 (1966).
5. R. J. Lasek and M. L. Shelanski, Cytoskeletons and the Architure of the Nervous System, Neuroscience Res. Prog. Bull. 19(1):79 (1981).
6. G. N. Phillips, E. E. Lattman, P. Cummins, K. Y. Lee and C. Cohen, Crystal Sructure and Molecular Interactions of Tropomyosin, Nature 278:413 (1979).
7. E. Lazarides and J. P. Revel, The Molecular Basis of Cell Movement, Scientific American 240(5):100 (1979).
8. J. S. Clegg, Intracellular Water, Metabolism, and Cellular Architecture, part II, in: "Coherent Excitations in Biological Systems," H. Fröhlich and F. Kremer, eds., Springer-Verlag, Berlin (1983).
9. D. E. Green, A Framework of Principles for the Unification of Bioenergetics, Ann. N. Y. Acad. Sci. 227:6 (1974).
10. J. L. Barker and B. R. Ransom, Pentobarbitone Pharmacology of Mammalian Central Neurons Grown in Tissue Culture, J. Physiol. 280:355 (1978).
11. A. Tzagoloff, "Mitochondria," Plenum, New York (1982).
12. D. E. Green and H. D. Vande Zande, Universal Energy Principle of Biological Systems and the Unity of Bioenergetics," Proc. Natl. Acad. Sci. 78:5344 (1981).
13. B. K. Siesjo, "Brain Energy Metabolism," John Wiley and Sons, Chichester (1978).

14. J. S. Richardson, The Anatomy and Taxonomy of Protein Structure, in: "Advances in Protein Chemistry," C. B. Anfinsen, J. T. Edsall and F. M. Richards, eds., Academic Press, New York (1981).
15. J. M. Hyman, D. W. McLaughlin and A. C. Scott, On Davydov's Alpha-Helix Solitons, Physica 3D 1&2:23 (1981).
16. A. C. Scott, Dynamics of Davydov solitons, Physical Review A 26:578 (1982).
17. A. T. Hagler, S. Lifson and E. Huler, The Amide Hydrogen Bond in Energy Functions for Peptides and Proteins, in: "Peptides, Polypeptides and Proteins," E. R. Blout, F. A. Bovey, M. Goodman and N. Lotan, eds., John Wiley, New York (1974).
18. J. F. Rabolt, W. H. Moore and S. Krimm, Vibrational Analysis of Peptides, Polypeptides, and Proteins.3.α-Poly (L-alanine), Macromolecules 10:1065 (1977).
19. N. B. Abbott and A. Elliott, Infra-red Spectrum and Dichroism of Crystalline Acetanilide, Proc. Roy. Soc. (London) 234A:247 (1956).
20. G. Careri, Search for Cooperate Phenomena in Hydrogen-Bonded Amide Structures, in: "Cooperative Phenomena," H. Haken and M. Wagner, eds., Springer-Verlag, Berlin (1973).
21. L. MacNeil and A. C. Scott, Launching a Davydov Soliton: II. Numerical Studies, Physica Scripta (in press).
22. V. A. Kuprievich and Z. G. Kudritskaya, Numerical Evaluation of the Exciton-Phonon Interaction Parameters in the Theory of Davydov Solitons in Polypeptide Chains, Reprint No. ITP-82-64E, Institute for Theoretical Physics, Kiev (1982).
23. J. Adriani, "The Chemistry and Physics of Anesthesia 2nd Ed.," Charles C. Thomas, Springfield (1962).
24. S. H. Roth, Differential Effects of Anesthesia on Neuronal Activity, in: "Molecular Mechanisms of Anesthesia, Progress in Anesthesiology Vol. 2," B. R. Fink, ed., Raven Press, New York (1980).
25. R. Buchet and C. Sandorfy, Perturbation of the Hydrogen-Bonded Equilibrium in Nucleic Bases. An Infrared Study, J. Phys. Chem. 87:275 (1983).
26. Y. Kyogoku, R. C. Lord and A. Rich, Specific Binding of Barbiturates to Adenine Derivatives, Nature 218:69 (1968).
27. H. J. R. Weintraub, Conformational Energy Calculations on Anticonvulsant Drugs, Int. J. Quan. Chem.: Quan. Biol. Sym. 4:111 (1977).
28. P. R. Andrews and J. A. Defina, Stereochemistry and Electronic Structure of Anticonvulsant Drugs, Int. J. Quan. Chem.: Quan. Biol. Sym. 7:297 (1980).

NONLINEAR PHENOMENA IN BIOENERGETICS AND ONCOLOGY AS SEEN IN 25 YEARS OF RESEARCH WITH MILLIMETER MICROWAVES AND RAMAN SPECTROSCOPY

Sydney J. Webb

Box 70. R.R.1. Fulford Harbour
B.C. Canada VOS 1CO

Studies over the past three or four decades into the chemistry of living entities have advanced conceptual views of life processes and have led to great progress in research technologies applied in every area of modern biology. It is understandable, therefore, that the predominant concepts of the nature of life are biochemical. As knowledge of the chemistry of living entities has grown, it has become clear that the vital reactions and syntheses inside a living cell take place in orderly sequences performed only at specific times in the lifetime of a cell. The existence of this metabolic, or cell timeclock is now well established. Also, observations of metabolism show in vivo rates of syntheses are many times faster than those observed in the test tube in which the reactants are isolated from the cell and its natural barriers. Such findings suggest that metabolites are directed to vital sites within the cell, with in vivo metabolism governed by a process able to direct the positioning of reactants in both space and time, rather than by random kinetics. A phenomenon such as this requires energy, and unless there is abundant waste, this energy must be directed to specific sites at which the reactions are to occur. Although it is known that the living cell is extremely efficient from the standpoint of its use of energy, it is not generally recognized that the cell does not use energy in the form of heat. Rather, heat often is a toxic waste of metabolism. In contrast to the high temperatures required in vitro to achieve the syntheses of biomolecules, such as amino acids, by purely chemical means, the living system performs the same synthesis at comparatively low temperatures by using the "configurational" energy present in large protein molecules in place of heat. Moreover, the cell operates by a process which must be considered isothermal, since changes of 4°C to 5°C above the optimum, can completely stop, or violently alter, in vivo chemical events. Although these facts are common knowledge today, bioenergetics still are discussed mainly in terms of units of heat, the calorie, and the heat released by the hydrolysis of phosphorylating agents such as ATP, even though much of this energy is never released as heat. In fact, much of the energy used to drive a given biological system is present in its enzymes and other giant macromolecules produced for it by a prior generation of cells (e.g. a mother cell). This energy is used to reorganize the vibrational or bond energies of supplied nutrients to form new macromolecules and hence, a daughter cell, with a minimal loss of energy in the form of heat.

It is clear that life is much more than simple test tube chemistry. It is a process involving the production of extreme order from randomness, directed energy flow and physical motions of molecules as well as a function of time. If therefore, the information provided by decades of work by biochemists, and by molecular and cell biologists is to improve conceptual biological ideas and biotechnology, then the knowledge from current research must also be integrated into new predictive models derived from a strongly interdisciplinary enterprise.

Viewed collectively, experimental data now available, indicate that the living cell is a unique assembly of macromolecules which acts as a single unit. Also, its ability to perform each of its many functions in a set time sequence, at rapid rates and at what must be considered as low temperatures suggests that it employs a form of electrodynamic property analogous to that of solid-state systems. Solid state physics with its accompanying mathematics and theories, may therefore be able to shed some light on the life process. The need for such an input into biological science is now more evident from modern experiments with microwaves and radiowaves.

We now turn to an historical account of work which has led to our present perspective. In the early 1950s, the author became interested in the relative stability of some plant viruses to stresses of dehydration and various kinds of radiation. This led to extensive studies over 15 years into the role of water molecules in the structure and function of in vivo macromolecules and complexes. As part of this study, the author calculated possible relaxation times of those water molecules confined to set spaces within macromolecules and bound to certain of their chemical groups. The results suggested that this water should absorb microwaves. In 1956 the author therefore began experiments with high frequency microwaves (50 to 150 GHz) to ascertain whether measured attenuation by cells held in aerosols at relative humidities (RH) of from 5% to 90% could yield new information concerning the amount and behavior of water bound to macromolecules in vivo at given RH levels. The initial series of experiments ascertained the desorption isotherms, at 33°C, as seen by the attenuation of set frequencies of microwaves. It soon became clear that the shape of the isotherm depended on the test frequency (Fig. 1). This implied that with cells held at 90% RH, where only one or two monolayers of water exist in vivo, their absorption spectra between 50 - 100 GHz should show peaks where attenuation by bound H_2O was strongest. The aqueous environment was minimized by subsequently holding cells at 90% RH in aerosols or in thin films (200 μ). Cells remained fully active and capable of division. This indicates that the natural environment of microbes on surfaces of various kinds and of animal tissues is imitated more closely by this film technique than with techniques in which cells are suspended in excess water. The technique also avoids obliteration of fine structure of microwave spectra due to gross and frequency independent attenuation by "free" water.

An aspect of these studies was the requirement that microwave fields in the range from 5 μW to 5 mW/cm^2, had no adverse effects on these cells. It soon became clear that while the microwaves had no observable action on dormant cells suspended in aerosols, they did affect growth characteristics if they were exposed during their recovery from the dormant state.

The unexpected finding of microwave effects reported in 1968 (1) prompted a study of bioeffects of these microwaves and their relationship to cell age and nutritional status, together with effects of field densities and frequency. This study is now in its 25th year, and while much has been learned it is only recently that the collected data has begun to yield a coherent picture of mechanisms. We will present selected data on salient features of these studies described below to illustrate certain points.

At this time, it seems that metabolic activity, especially the operation of the respiratory chain, is prerequisite for the reception of the "non-thermal" response to microwaves (2); the cell, therefore, like many crystals has to be "activated" in order for it to respond to these waves. In its active state, moreover, the cell can not only receive such energy but amplify it. Thus nonthermal effects of microwaves can alter metabolic activities, and, indeed, the vital time sequences in which they are performed. The activation of the membrane respiratory chain seems to be required to facilitate reception of this energy yet the effects seen are on the synthesis of DNA and its coding of many proteins (2). Thus, the radiant energy appears to be either absorbed by the membrane and then transferred to the DNA or alternately metabolically produced energy in the membrane is transferred to the DNA and this renders it receptive to the incoming photons. In either case, however, energy is moved from one part of the cell to another. More important still are the facts that the cell only responds to certain frequencies of such waves and the particular frequencies to

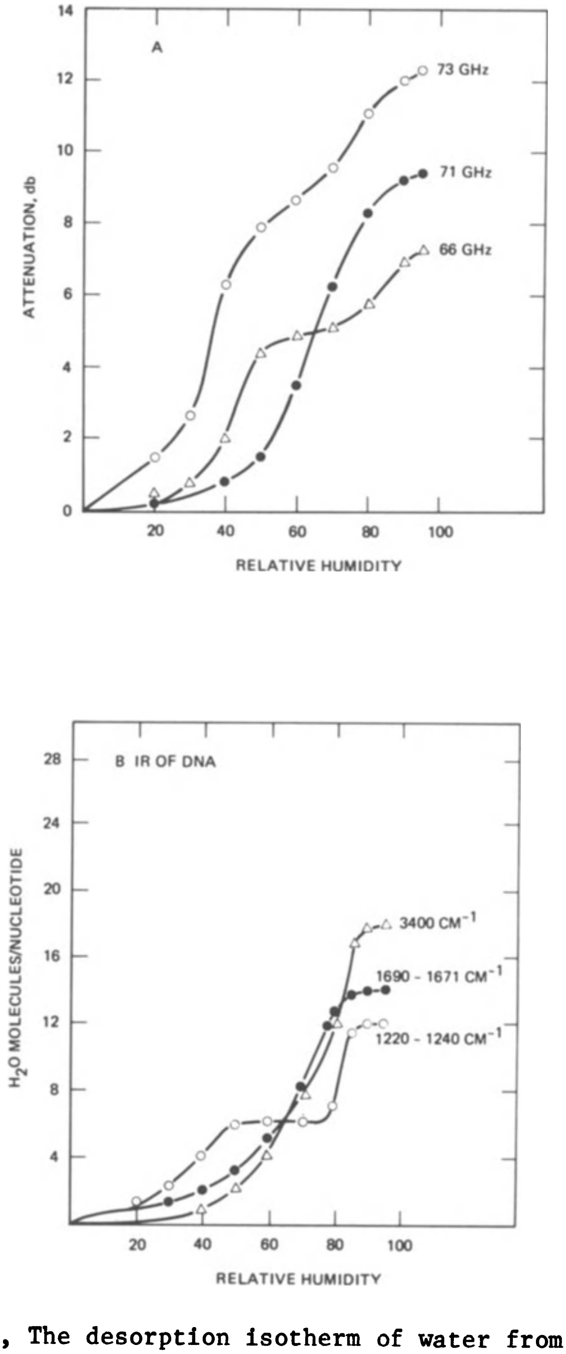


Fig. 1 (A), The desorption isotherm of water from cells of E. coli at three different frequencies of microwaves: (B), The number of water molecules bound to the groups of DNA at different relative humidities as seen by infrared spectroscopy. Taken from Webb, 1976 (5).

which it will respond alter with its age and nutrition (2, 4, 5)(Table 1).

Thus both the stage the cell has reached during its lifetime and its nutrition, especially the type of carbon supplied (e.g. oxidizable energy source), determine which frequencies the cell is able to receive.

One of the features of the data collected, shown in Table 1 was that the effective frequencies formed a definite nonlinear series which suggested that they all were related via the "harmonics" of some fundamental oscillatory in vivo mode or modes. This became important later when it was realized that although the actual effective frequencies changed with time and nutrition their nonlinear relationship did not, in every case two definite series could be clearly demonstrated. When we compared the spectra of mouse mammary carcinoma cells with their homologous normal ones, we found that the effective frequencies for normal cells formed into the usual two nonlinear series, while those of tumor cells formed into three and sometimes four series. We observed the same splitting of a single series into two or three separate ones later when bacterial cells were infected with a bacteriophage, BHK cells with Mouse Sarcoma Virus (MSV) (Fig. 2), or when we compared biopsies of human mammary carcinoma tissue to normal mammary tissue. One interesting feature of the latter study was that the "normal" tissue from the undiseased mammary gland of the patient with carcinoma of the other also showed a single split in one of the normal series of nonlinear harmonics wheras the diseased tissue from the other mammary gland showed a double split, i.e. four distinct series (Fig. 3). It appeared, therefore, that in the "normal" tissue a single split had occurred whereas in the diseased tissue a double split had taken place. This finding led the author to suggest that an examination of mammary tissue by microwave spectroscopy might provide a method by which the early presence of, or even predispositon to, mammary carcinoma could be detected. Although the data presented in Figs. 2 and 3 are ratios of energy absorbed in cells to energy absorbed in the medium, each set at each exposure frequency had a proportional influence on the uptake and incorporation of ^{14}C aminoacids and nucleotides. Therefore, the spectra are both absorption and action spectra. Absorption ratios are shown because the bioeffect did not always manifest itself in an inhibition, in some cases a stimulation was observed and this made it more difficult to present the data graphically. Whenever the ratio between attenuation by the medium alone and the cells plus medium was unity, no bioeffect was ever seen.

Recently Raman spectroscopy has been employed in attempts to display the internal oscillations of whole cells. It was considered at the time this approach was conceived that a very large number of oscillations would be present and these would produce a

TABLE 1

The Effect of Microwaves on the *In Vivo* Biosynthesis of Macromolecules

(A) CELLS IN GLUCOSE - AA

Frequency (GHz)	59	61	66	71	73	76	126	129	136	141	143
Cell growth	0.11*	0.52	0.09	0.51	0.10	0.49	0.53	0.08	0.11	0.51	0.08
DNA synthesis	0.08	0.46	0.07	0.49	0.12	0.52	0.48	0.11	0.09	0.53	0.10
RNA synthesis	0.97	0.98	1.02	1.10	0.84	0.99	1.02	1.04	0.98	0.99	0.79
Protein synthesis	0.09	0.47	0.08	0.51	0.11	0.56	0.48	0.10	0.12	0.51	0.09

(B) CELLS IN GLUCOSE-NH_4^+

Frequency (GHz)	59	61	66	71	73	76	126	129	136	141	143
Cell growth	0.54	0.08	0.10	0.11	0.48	0.06	0.09	0.49	0.06	0.11	0.52
DNA synthesis	0.48	0.10	0.08	0.12	0.46	0.12	0.10	0.48	0.08	0.10	0.48
RNA synthesis	1.01	1.11	1.08	0.84	0.84	0.96	1.06	1.04	0.92	0.98	0.81
Protein synthesis	0.51	0.09	0.11	0.07	0.51	0.11	0.13	0.52	0.10	9.09	0.47

(C) FREQUENCIES ABLE TO AFFECT RNA SYNTHESIS

(strong)	63	68	73	78	123	128	133	
(weak)	70	76	79	124	127	130	136	139

---

*Results expressed as fractions of the amounts synthesized by control non-irradiated cells.
Taken from Webb, S.J. (1975) (2)

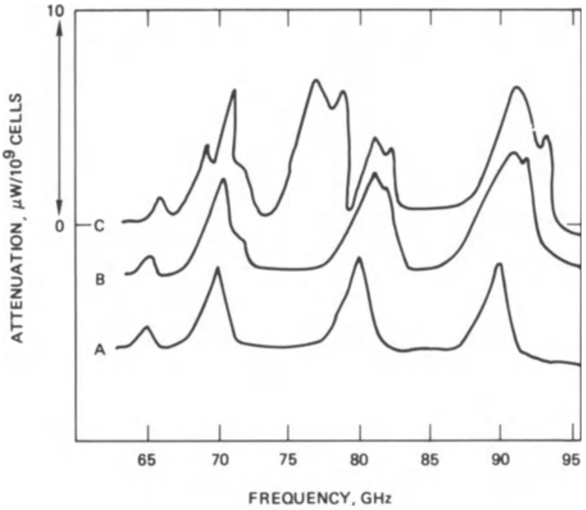


Fig. 2(a) Microwave spectra. A - BHK cells. B - MSV infected BHK cells at 48 hours. C - MSV infected BHK cells at 72 hours.

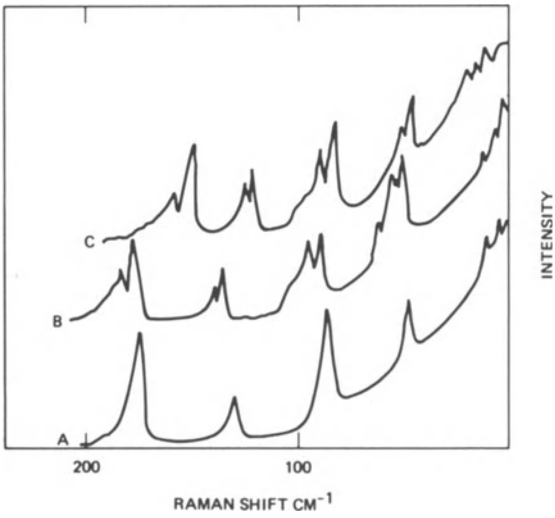


Fig. 2(b) Laser-Raman spectra. A - BHK cells. B - MSV infected BHK cells at 48 hours. C - MSV infected BHK cells at 72 hours. Taken from Webb, et al., 1977 (15, 16).

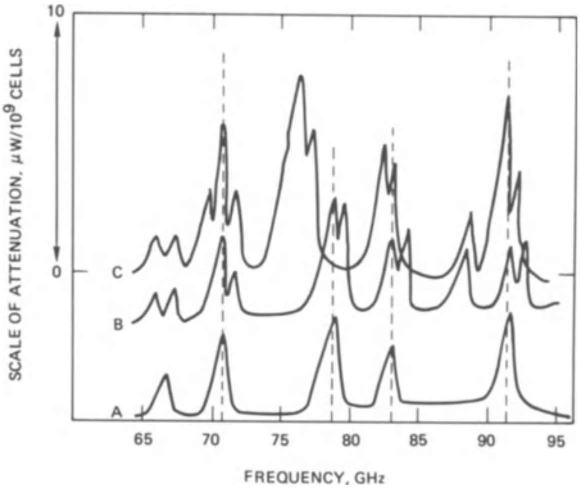


Fig. 3(a) Microwave spectra. A - normal mammary tissue. B - "normal" right-side tissue. C - left-side 2-3 cm diameter mammary carcinoma. Spectra are presented as the ratio of attenuation by cells in the medium over that of the medium alone.

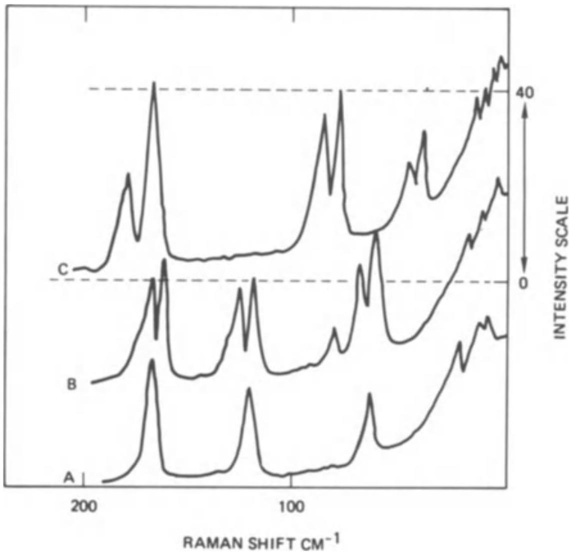
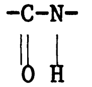


Fig. 3(b) Laser-Raman spectra. A - normal mammary tissue. B - "normal" right-side tissue. C - left-side 2-3 cm diameter mammary carcinoma. From Webb, et al., 1977 (15, 16).

very complex spectrum of lines. However, no spectrum was found with resting, living cells (6). Once the cell was activated by the addition of a suitable nutrition containing a usable oxidizable carbon source, a Raman spectrum appeared, but the spectrum changed continuously as the cell proceeded through its life cycle (Fig. 4). Later when these spectra were analyzed with respect to time, sets of lines between 200 and 3400 cm^{-1} were found to move gradually to higher frequencies while those between 5 and 200 cm^{-1} moved to lower ones (Fig. 5). Another interesting finding was that just before cell division the spectrum became essentially devoid of lines except for one or two lines of high intensity and at high frequencies, around 2100 cm^{-1}, which suggested that a large portion of available energy at this stage was devoted to the division process (Fig. 6). The lines above 200 cm^{-1} have been ascribed to oscillations in large complexes of macromolecules excited by metabolically induced energy, a possibility predicted on theoretical grounds by Frohlich (7). An indication of the occurrence of condensations and subsequent coherent oscillations is shown in Fig. 7. This pattern is one of the most reproducible spectral events in active bacterial cells and is seen in almost all spectra at a set time in the life of the cell. (A) is the spectrum at the start of the cell cycle; (B) at division; (C) at 20 minutes and (D) at 40 minutes respectively through an 80 minute cell lifetime. At division two well defined lines appear at 820 and 2100 cm^{-1} respectively. These then decrease in intensity, while the broad line at 1640 cm^{-1} becomes a little narrower in width with a high intensity narrow peak. The lines between 1620 to 1690 cm^{-1} are produced by vibrational stretching nodes of the -C=N- of the ring structures of nucleic acids, the  $\underset{\displaystyle O}{\overset{\|}{-C}}\underset{\displaystyle H}{\overset{|}{-N-}}$ of the peptide bond and the bending of water molecules. We assume that Raman lines in this region arise from the same vibrations. If so, the spectral changes seen with time may represent a condensation and induced common oscillation between these three components (e.g. DNA (or RNA) + protein + H_2O). A recent mathematical analyses of the lines above 200 cm^{-1} has shown that they all emanate from 2 to 4 fundamental modes, that is, all lines appear to be the nonlinear "harmonies" of 2 to 4 fundamental *in vivo* oscillations (8).

The Raman lines below 200 cm^{-1} (examples in Fig. 8) persist for much longer periods during the cell's lifetime and moving slowly to lower energies. Scott (9) has suggested that they correspond to the fundamentals and harmonics of a solitary wave, known as a soliton, traveling through large macromolecules and complexes of them. The frequencies and hence energy, of such a wave were calculated from theory and the known structure of proteins without knowledge of the Raman data; it was gratifying therefore, that the theoretical and experimental data fitted so well (10). The soliton as described by Davydov (11), is a wave

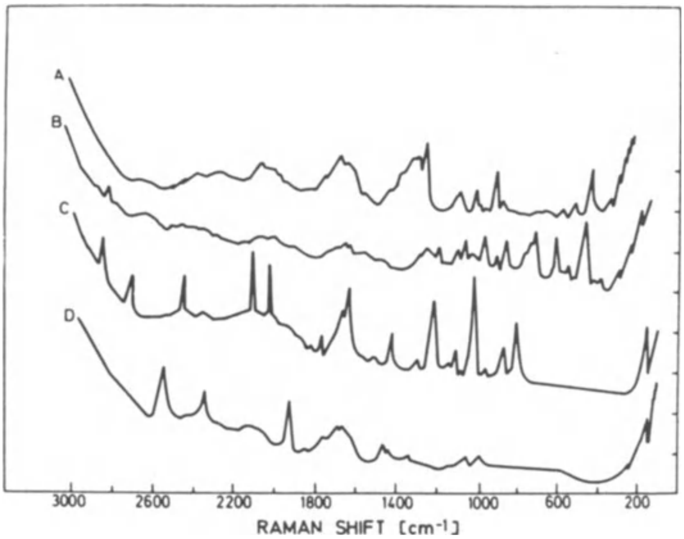


Fig. 4(a) Examples of spectra obtained with synchronized Escherichia coli cells in Davis NH_4^+ salts. A - 0-20 min. B - 20-40 min. C - 40-60 min. D. 65-85 min. Conditions as for Fig.1.

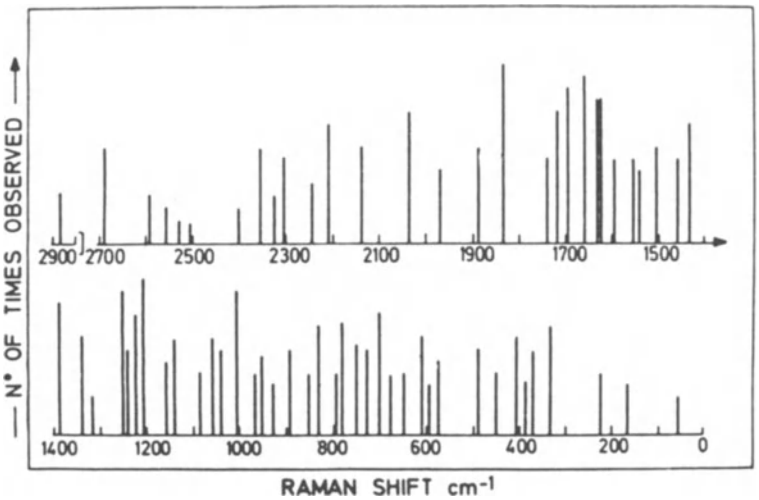


Fig. 4(b) Composite of all Raman lines observed in the spectra obtained with synchronized populations of Escherichia coli cells. Each Raman line seen in a given spectrum, recorded at a set time in the lifetime of the cell, is tabulated along with the number of times each line re-occurred in repeat spectral determinations. The height of the frequency markers, used to scale, the illustration = two appearances of a line. From Webb, 1980 (6).

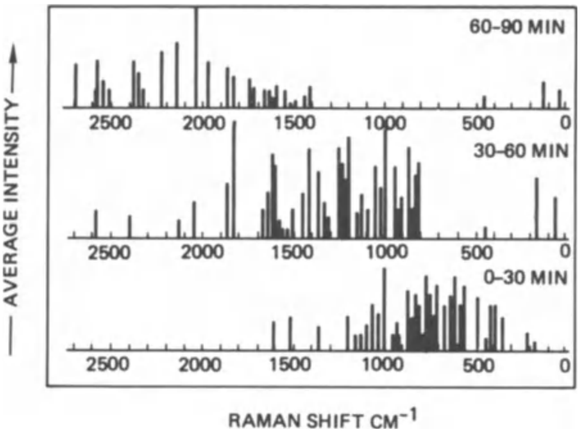


Fig. 5 The timed appearance of Raman lines seen in the spectra of synchronized populations of Escherichia coli cells during a 90 minute cell cycle. The times shown represent the time after the cells were placed in a minimal medium and, at which, their Raman shift spectrum was determined. The average intensity of each line is that calculated from two or more spectra in which the line appeared and is presented merely to demonstrate the relative intensities of the lines and change in them with time. From Webb, 1980 (6).

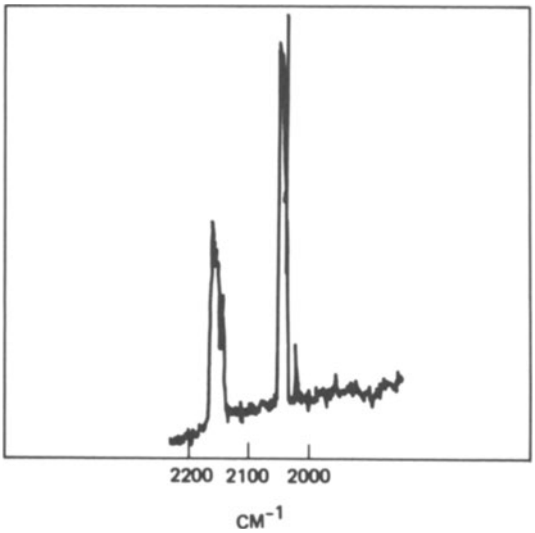


Fig. 6 The strong Raman lines which appeared in the spectrum of Escherichia coli just before cell divisions began. Laser 5145 A, sensitivity 10 K. Taken from Webb, 1980 (6).

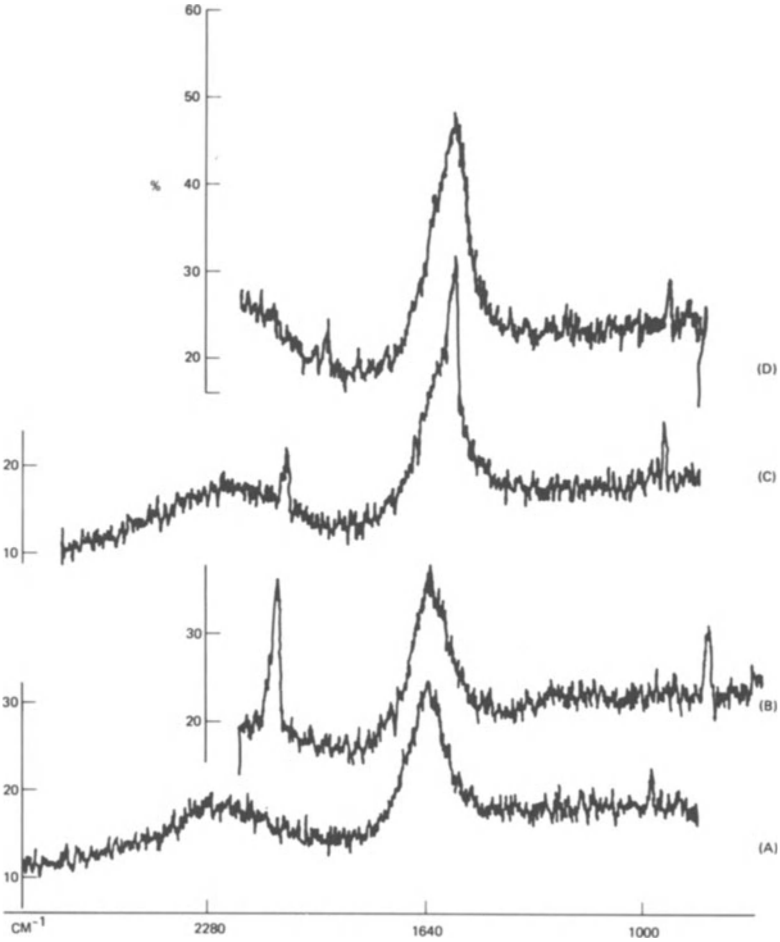


Fig. 7. Raman spectra of synchronized cells of *Escherichia coli* (see text).

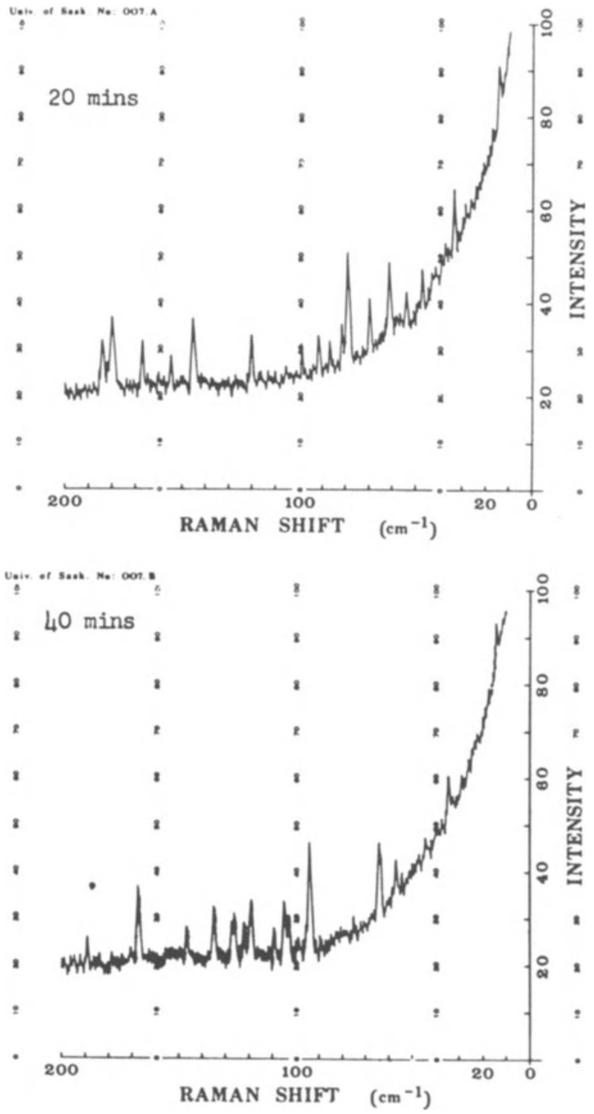


Fig. 8. Low frequency Raman spectra of synchronized Escherichia coli cells taken from Lomdahl, et al., 1982 (10).

form which does not dissipate, but moves along a protein alpha-helix until all of its energy is dissipated at some point at the end. From the biological point of view it may be considered as a localized conformational change which then travels down or through a chain of molecules. These waves are characterized by two fundamental frequencies, one arising from the oscillation that produces the wave, and the other from the velocity with which the wave travels. The process is a nonlinear one and so other frequencies arise from these fundamentals by harmonic generation and frequency mixing.

The initial good fit between experimental and the theoretical soliton was less obvious later when the frequencies of the lines in a large number of Raman spectra were added in the form of a histogram. This loss of fit seems largely due to the shift that occurs in the frequencies of the line with time as the cell ages. The author therefore constructed a histogram of lines in eight spectra taken at the same time in the life cycle of *Escherichia coli* (Fig. 9). The fit with the earlier theoretical fundaments of 23 cm^{-1} and 126 cm^{-1} for the frequencies of the oscillation and wave velocity respectively in proteins, is good, but it is of interest that a fundamental of 16.2 cm^{-1} is also present. A theoretical analysis was made with the assumption that two types of soliton existed *in vivo*, one in protein and the other in DNA as these are the two major types of biomolecule able to act as media for the propagation of such waves. The analysis also assumed that the velocity of the waves was the same having a frequency of 126 cm^{-1}. By nonlinear analysis, the "harmonics" of each wave was calculated and it was found that the frequencies of some harmonics within reasonable limits, were common to both waves.

The above experiments with bacterial cells were repeated with similar studies with normal and tumor cells. The low frequency Raman spectra (5 to 200 cm^{-1}) of normal cells always displayed a series of sharp single lines which formed two nonlinear series of "harmonics", as in the microwave data. In spectra of all types of tumor cells studied, however, these lines were broader, of lower intensity, and split into two or three separate lines. In an attempt to understand this phenomenon the Raman spectrum of normal BHK cells was established and changes subsequent to infection with MSV was followed over two or three days. The spectrum of infected cells changed gradually after infection with early splitting of lines and more pronounced changes at later times. When no evidence of the virus could be detected in the culture medium, after several subcultures of infected cells, the spectrum displayed split broad lines (see Figs. 2 & 3). Finally when the microwave and Raman data were viewed collectively, it became clear that the low frequency Raman lines reflected the upper nonlinear "harmonics" of two fundamental oscillations which resided somewhere in the millimeter microwave region. The question then is can this process be

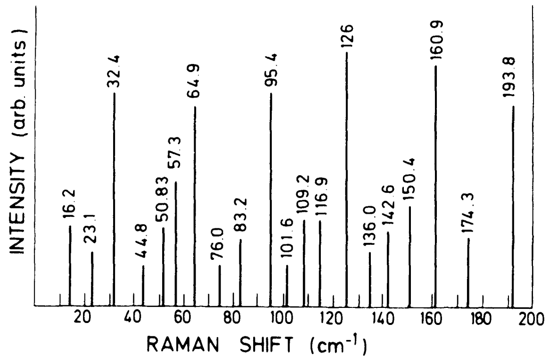


Fig. 9. Composite of low frequency Raman lines of *Escherichia coli* at 30 minutes.

reversed and the normal oscillations restored? It is likely that this could be done by chemical means but the steriochemistry of any effective agent will differ from tumor to tumor and a specific one may be needed for each individual tumor. A much more general approach would be to expose the tumor to two close microwave frequencies it absorbs, and then modulate between these two at a frequency characteristic of the normal cell. This has been tried by the author with a modicum of success (2, 14) (Fig. 10).

The question is what can now be learned or implied, from this new type of biophysical data? The Fröhlich coherent oscillations of macromolecules provides possible mechanisms by which the living cell is able to receive and respond to microwaves of specific frequencies. Second, the concept of energy transfer by solitons offers a means by which energy may move rapidly along set pathways to specific sites in the cell. A system of linked filaments of protein which grows in a unidirectional way with time has been observed in living cells (12, 13). Such a network could act as a vector in vivo for solitary waves. Third, the existence of possible interference between the traveling waves could be one mechanism by which a given metabolic activity is confined to a given place and time. Sufficient energy could be delivered to drive a reaction only at junctions of the wave tracks. The energy of the line at 2100 cm^{-1} and its amplitude suggests that considerable energy is being delivered to one or two places in the cell to effect cell division. It seems plausible to suggest that the spindle apparatus of eucaryotic cells, being protein, could act as a vector of this energy, via solitons.

Evidence that a cell may generate more than one fundamental frequency, and also produce their harmonics as a manifestation of metabolically induced coherent oscillations in large macromolecular complexes; and evidence for the transfer of this "oscillatory" energy as solitary waves along set pathways to specific sites in the cell suggest mechanisms by which in vivo reactions may be controlled in time and place, rapidly performed and kept "in tune" with one another. It is well known that in proliferating cells, in a stable nutritional environment, the ratio between the amounts of three major components, protein, RNA, and DNA, remain constant. Only when the nutrition is changed does this ratio change and then only in the generation of cells experiencing the change. This constant ratio can only be maintained if the respective rates of syntheses of these macromolecules exhibit the same ratios relative to one another. Therefore, these rates are closely controlled.

The nonlinear approach to biological energy transactions, and the concept of the cell in terms of its oscillations rather than its chemical activities provide new models of living processes and of neoplastic changes, concepts of more than mere theoretical interest. Several promising lines of investigation of nonlinear millimeter wave effects have been reviewed in these proceedings. The present prospect for new knowledge is exciting.

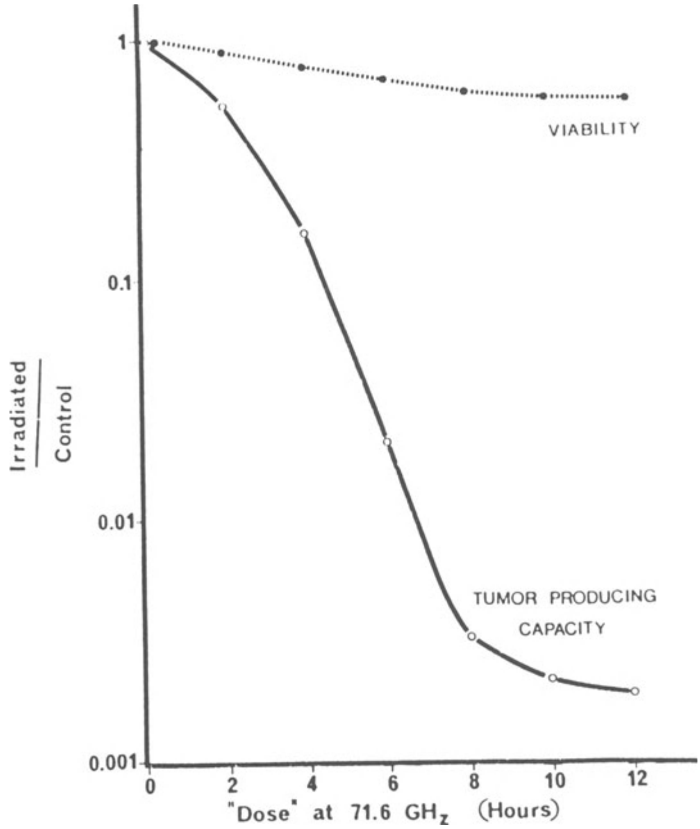


Fig. 10. Effect of exposure to microwaves on the ability of virus-transformed BHK tumor cells to form clones in tissue culture and produce tumors in experimental animals. Inoculum per animal, 10^7 cells. Taken from Webb, 1975, (14).

REFERENCES

1. Webb, S.J. and Dodds, D.E. *Nature* *218*:374 (1968).
2. Webb, S.J. *Annals N.Y. Acad. Sci.* *247*:327-352 (1975).
3. Webb, S.J. IRCS *J. Med. Sci.* *4*:10 (1976).
4. Webb, S.J. IRCS *J. Biochem-Biophys.* *2*:1232 (1974).
5. Webb, S.J. In: "Nutrition, Time and Motion in Metabolism and Genetics", C.C. Thomas Pub., pp. 246-275 (1976).
6. Webb, S.J. *Physics Reports* *60*:201-224 (1980).
7. Frohlich, H. Advances in Electronics and Electron Physics *53*:85-152 (1980).
8. Del Giudice, E., Doglia, S. and Milani, M. *Phys. Lett.* *85A*:402-404 (1981).
9. Scott, A.C. *Phys. Lett.* *86A*:60-62 (1981).
10. Lomdahl, P.S. et al., *Phys. Lett.* *92A*:207-210 (1982).
11. Davydov, A.S. *In*: "Biology and Quantum Mechanics", Pergamon, Oxford (1982).
12. Wolosewick, I.I. and Porter, K.R. *J. Cell. Biol.* *82*:114-120 (1979).
13. Hyams, J. *Nature* 291:107 (1981).
14. Webb, S.J. *Int. J. Quant. Chem. Quant. Biol. Symp.* *1*:245-251 (1975).
15. Webb, S.J. and Lee, R. IRCS *Med. Sci.* *5*:102 (1977).
16. Webb, S.J., Lee, R. and Stoneham, M.E. *Int. J. Quant. Chem. Quant. Biol. Symp.* *4*:277-284 (1977).

NONLINEAR ELECTRODYNAMICS IN CYTOSKELETAL PROTEIN LATTICES

Stuart R. Hameroff
Department of Anesthesiology
University of Arizona Health Sciences Center
Tucson, Arizona

Steven A. Smith
Computing Division
Los Alamos National Laboratories
Los Alamos, New Mexico

Richard C. Watt
Department of Anesthesiology
University of Arizona Health Sciences Center
Tucson, Arizona

INTRODUCTION

Nonlinear electrodynamic theories predict dynamic organization of biomolecular activities at all cellular levels: DNA, membranes, extracellular glycoproteins, and the interconnecting cytoskeletal lattice.[1-5] Cytoskeletal lattice proteins including microtubules are particularly involved in dynamic regulation of intracellular movements and activities.[6,7] This paper considers possibilities and implications of biological information processing due to coupling of Davydov solitons, Frohlich coherent oscillations and other nonlinear electrodynamic phenomena to conformational states of the grid-like polymer subunits of cytoskeletal microtubules.

CYTOSKELETAL PROTEIN LATTICES

Cytoskeletal lattices (Figure 1) including protein polymers microtubules (MT), actin, intermediate filaments and the microtrabecular lattice interconnect and organize cell components and dynamic

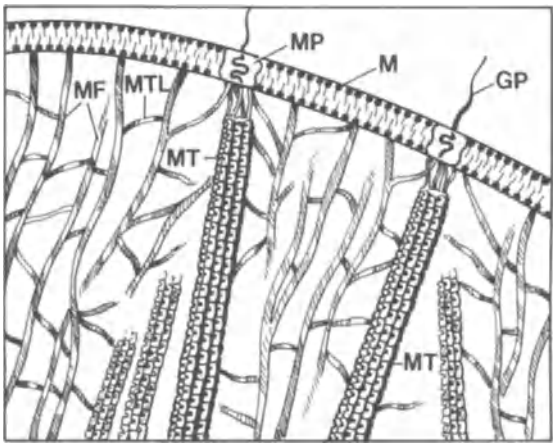


Fig. 1. Schematic representation of cytoskeletal lattice elements in vicinity of cell membrane. MT: microtubule, M: membrane, MP: membrane protein, MF: microfilaments, MTL: microtrabecular lattice, GP: extracellular glycoprotein.

activities.$^{5-7}$ Intracellular movement and coordination, growth, mitosis, transport and turnover of membrane proteins are accomplished by activities of cytoskeletal lattices and attachments. The major cellular architectural elements, MT are cylindrical polymers which also comprise cilia, mitotic spindles, and other organelles.$^{8-11}$ MT are intimately involved in dynamic biological activities, but mechanisms of "real time" regulation and control by MT or other cytoskeletal filaments are unknown. The particular structural geometry of MT will be considered as a potential site for functionally significant nonlinear electrodynamic phenomena.

Most MT are assemblies of 13 longitudinal protofilaments which are each a series of polar, tubulin dimers (Figure 2). Each dimer consists of two slightly different classes of 55,000 molecular weight monomers (α and β tubulin) which are 4 nm in diameter.$^{12-14}$ Tubulin binds eighteen calcium ions (Ca^{++}) per dimer, and MT polymerization depends on Ca^{++} and other factors. Hexagonal packing of the dimers

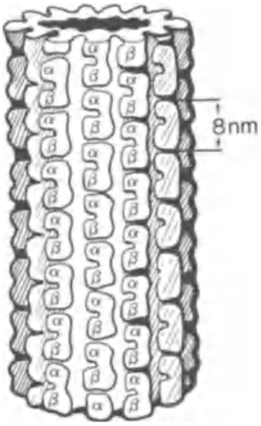


Fig. 2. Microtubule (MT) structure from x-ray crystallographic data.[12] MT subunits are 8 nm dimers of α and β tubulin in a hexagonal, cylindrical lattice.

results in 3, 5 and 8 pitch helical patterns.[15,16] MT outer diameters are 25-30 nm and inner diameters are 14-15 nm.[17,18] MT outer surfaces are surrounded by a "clear zone" of several nm which apparently represents electronegative charges on tubulin[19] and may also organize cytoplasmic water and enzymes.[5,6] Functional MT lengths may range from hundreds of nm to micrometers and meters in some mammalian neurons. Contractile (ATPase dynein) or enzymatic proteins (microtubule associated proteins: "MAPs") may be attached at specific dimer sites resulting in various helical patterns.[20] Numerous proteins and other substances are transported by MT and attached contractile proteins at rates from one to 400 millimeters/day.[21-23] Phosphate nucleotides including GTP and cyclic GMP bind reversibly at specific dimer sites.[24-25]

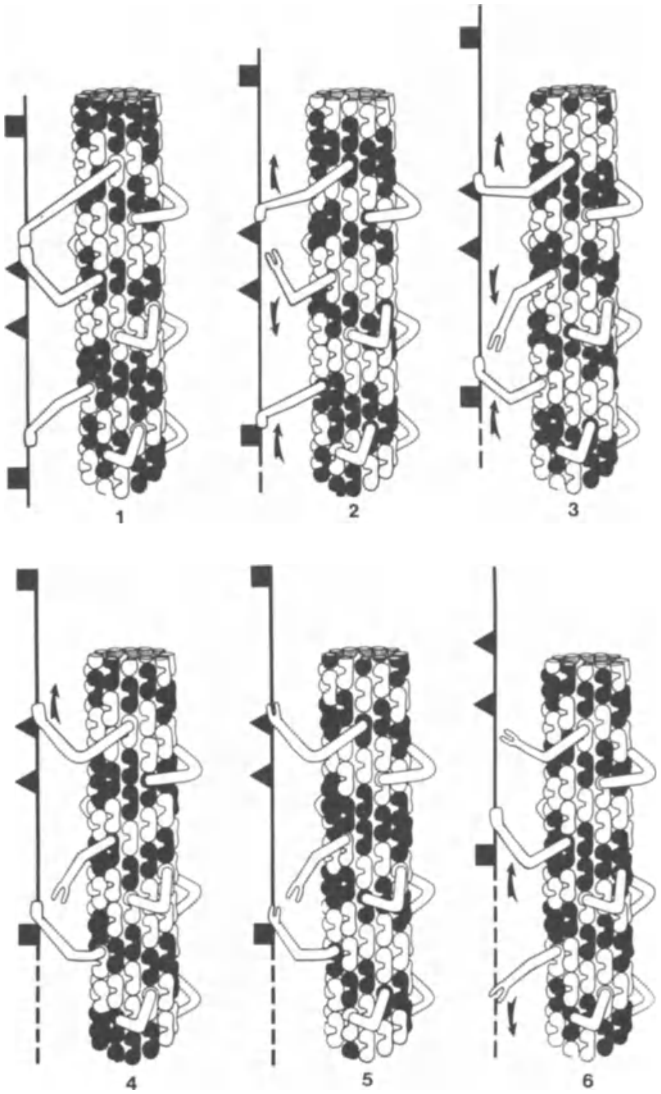


Fig. 3. Coordinated transport by MT based ATPase contractile sidearm proteins. Signaling for contraction may occur by solitons propagating through the MT lattice. In this representation, inclusion of sidearm base dimers in a "diamond" (4 contiguous aligned dipoles) connotes signal to contract. Diamonds or other structures may derive from coherent oscillations (see automaton model) resulting in solitons.

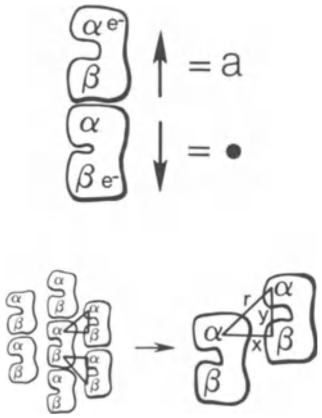


Fig. 4. Dimer subunits in automaton neighborhood. Assuming α and β monomers share an electron, dimer is in α state (represented by a) if dipole orientation is toward α monomer, and dimer is in β state (represented by •) if dipole orientation is toward the β monomer. Relative distances among neighbor loci may be represented by r, x, and y.

Extensive reviews on MT and cytoskeletal structure and function are available.[8,11] Intimate involvement of MT with dynamic cellular activities and their grid-like polymer structure suggest capabilities for information processing which may be explained by nonlinear electrodynamic phenomena.

SOLITONS, COHERENT OSCILLATIONS, AND DYNAMIC ORGANIZATION OF PROTEIN CONFORMATIONAL STATES

Solitons

Biological solitons have been theorized as couplings of phonons and excitons along molecular chains producing nonlinear vibrations which are extremely long-lived in comparison to linear oscillations.[1] Davydov proposed that biological solitons could be energetically initiated by hydrolysis of ATP phosphate bonds and propagate along linear molecules such as the alpha helix portion of proteins.[2,4] Nonlinear coupling of electronic bond energy (possibly anharmonic stretching of amide 1 bonds,[4,26] orbital changes in hydrophobic aromatic regions, and/or disturbance of the electronegative charge field on the MT outer surfaces) or mechanical oscillations of protein peptide chains can account for wave-like soliton propagation and a control mechanism for protein conformational states. Calculations and simulations have suggested that solitons may propagate through an alpha helix protein sequence at subsonic velocity, occupying nine peptide bond lengths (∿4 nm) per soliton.[4,26-28]

Cytoskeletal lattice arrays may provide significant pathways for soliton propagation. Phosphate nucleotides such as GTP bind rapidly and reversibly at specific tubulin monomer sites. GDP, incapable of energy donation, binds tightly at other tubulin monomer sites.[29,30] Thus solitons may be introduced or supported at specific "preprogrammed" GTP binding subunits. Tubulin monomers are 4 nm in diameter and are comprised of about 22 percent alpha helix chains.[14,31] Thus solitons initiated by GTP and other factors could propagate between α and β tubulin dimer regions, altering conformational states and binding of Ca^{++} and other factors. Soliton propagation among neighboring MT grid dimers might depend on numerous additional factors including dimer-dimer neighbor interactions, dimer conformational and Ca^{++} binding states, attachment of GTP, MAPs and contractile proteins, and perhaps the directional orientation of alpha helix chains within the globular tubulin subunits. Solitons propagating through an MT lattice would encounter gates and switches at every dimer-dimer interface. Dynamic patterns of solitons propagating and switching through MT lattices have been compared to computer and information processing technologies (Boolean switching matrices, charge coupled devices, bubble memories, etc.) and theoretically linked to dynamic organization of cellular activities.[6] Programming modes governing soliton information patterns within cytoskeletal lattices may include genetic (e.g., primary protein structure) and "environmental" (e.g., Ca^{++} flux) influences.

Biological activity which may be explained by soliton propagation and switching in MT lattices is exemplified by the possible orchestration of contractile proteins attached at specifically arrayed MT lattice sites and which are active in transport of mole-

Table 1. Automaton rules derived from $\Sigma\ y/r^3$ which may be calculated from MT neighbor geometry distances.

Neighbor	Center State	Neighbor State	Distance (x)	Distance (y)	y/r^3
N	alpha	alpha	0.	8.	.0156
N	alpha	beta	0.	4.	.0625
N	beta	alpha	0.	12.	.0069
N	beta	beta	0.	8.	.0156
NW	alpha	alpha	-5.	4.9231	.0314
NW	alpha	beta	-5.	0.9231	.0064
NW	beta	alpha	-5.	8.9231	.0083
NW	beta	beta	-5.	4.9231	.0314
SW	alpha	alpha	5.	-3.1769	-.0152
SW	alpha	beta	5.	-7.1769	-.0108
SW	beta	alpha	5.	0.9231	.0070
SW	beta	beta	5.	-3.1769	-.0152
S	alpha	alpha	0.	-8.	-.0156
S	alpha	beta	0.	-12.	-.0069
S	beta	alpha	0.	-4.	-.0625
S	beta	beta	0.	-8.	-.0156
SE	alpha	alpha	5.	-4.9231	-.0314
SE	alpha	beta	5.	-0.9231	-.0083
SE	beta	alpha	5.	-8.9231	-.0064
SE	beta	beta	5.	-4.9231	-.0314
NE	alpha	alpha	5.	3.1769	.0152
NE	alpha	beta	5.	-0.9231	-.0070
NE	beta	alpha	5.	7.1769	.0108
NE	beta	beta	5.	3.1769	.0152

cules and organelles. Figure 3 depicts such orchestration by the signaling of propagating solitons. Dynamic organization of intracellular transport, ciliary movement, Ca^{++} fluxes, patterns of cell growth, differentiation and other intricate biomolecular actions requiring temporospatial organization could thus be explained.

Coherent Dipole Oscillations

Frohlich has theorized that nonlinear dipole excitations among molecular subunits of membranes and proteins of high dielectric

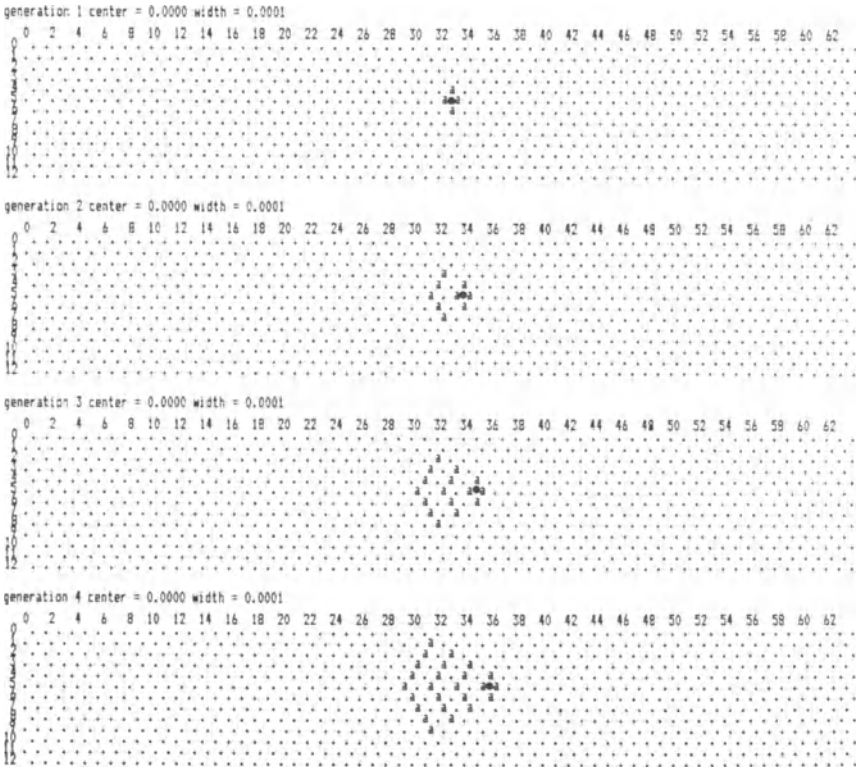

generation 1 center = 0.0000 width = 0.0001
generation 2 center = 0.0000 width = 0.0001
generation 3 center = 0.0000 width = 0.0001
generation 4 center = 0.0000 width = 0.0001
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62

Fig. 5. Computer simulation of MT automaton based on rules derived from electrostatic force contributions from neighbors. Protofilaments 0-12 are on vertical (x) axis, and rows 0-64 are on horizontal (y) axis. An all beta (') pattern is invaded by an initial pattern of alpha states which grows. Diamond patterns of alpha states are highlighted by a larger central dot. Generations 1-4 are depicted.

strength will result in coherent oscillations (10^{-10} to 10^{-11} sec fluctuations), lower frequency metastable states, coherent polarization waves ("solitons"), and long-range order.[2,32-34] Coherent oscillations of MT and other cytoskeletal lattice subunits may derive from intrinsic oscillations dependent on a high dipole moment and dielectric strength across the MT wall[19,35] or coupling to oscillations in membranes or other structures. Del Giudice et al.[35,36] have suggested that MT and other cytoskeletal filaments entrain and focus depolarization waves as solitons by nonlinear waveguide behavior.

Frohlich also theorized that protein conformational states are coupled to electron dipole oscillations.[33] Protein conformational

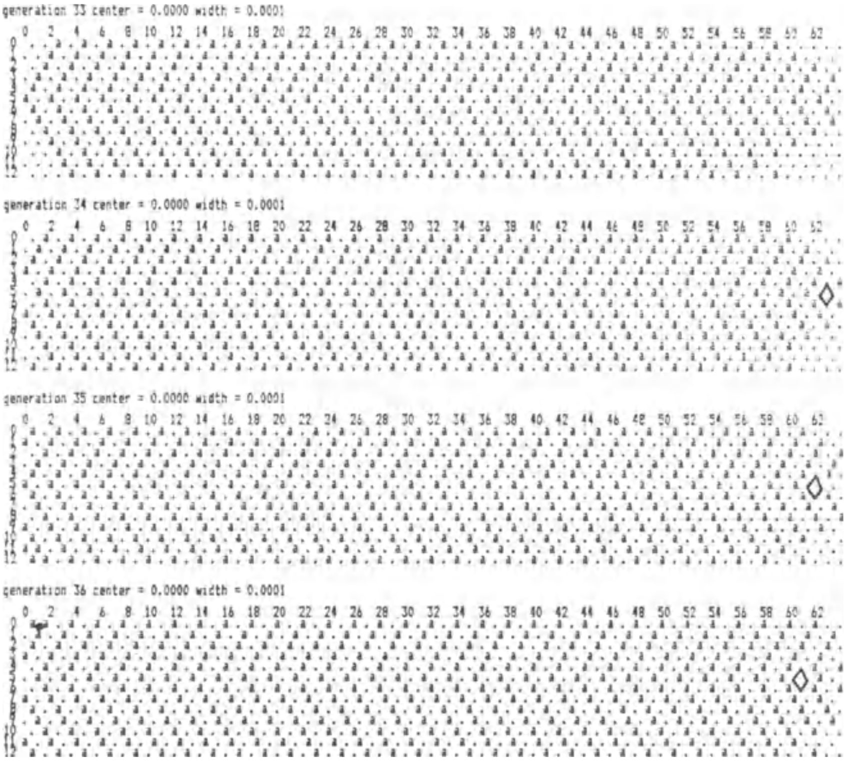

generation 33 center = 0.0000 width = 0.0001
generation 34 center = 0.0000 width = 0.0001
generation 35 center = 0.0000 width = 0.0001
generation 36 center = 0.0000 width = 0.0001

Fig. 6. Generations 33-36 of the patterns from Fig. 1. The alpha states have grown and reached the boundaries causing alpha diamonds and beta diamonds, highlighted by line connections. The pattern is otherwise orderly and alternating alpha and beta.

coupling to coherent dipole excitation among individual subunits in biomolecular arrays conceptually resembles soliton propagation and would apparently depend on genetic and environmental biochemical factors as well as neighbor interactions. Protein polymer neighbor interactions can be mediated by factors which include electrostatic dipole interactions, conformational states, binding of water, ions, MAPs, phosphate nucleotides and associated proteins, intrinsic subunit factors, or solitons. Net effects of these influences could alter oscillatory phases of particular subunits in coherent biomolecular arrays, possibly resulting in dynamic patterns, polarization waves and long-range order. Dynamic and stable patterns of dipole excitation and conformational states in MT could have direct, profound effects on regulation and organization of cellular

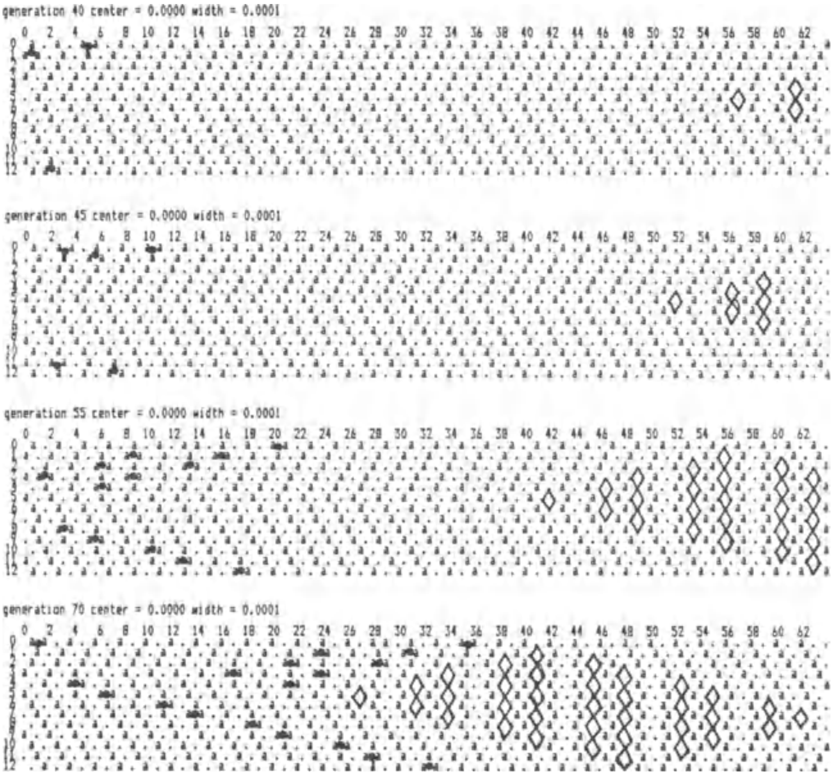

generation 40 center = 0.0000 width = 0.0001
generation 45 center = 0.0000 width = 0.0001
generation 55 center = 0.0000 width = 0.0001
generation 70 center = 0.0000 width = 0.0001
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62

Fig. 7. Generations 40, 45, 55, and 70 subsequent to Fig's 5 and 6. Alpha and beta diamond patterns develop further from boundary interactions and move towards each other at velocities of 8 nm/10^{-10} sec (∿80 m/sec).

processes. As proposed for soliton traffic, patterns of resonant excitation and conformational states of cytoskeletal lattice subunits could mediate protein binding sites, ion fluxes, and organization of cellular activities. Self-organizing patterns and information processing based on neighbor interactions at discrete time intervals among lattice subunit states have been described and are defined as cellular automata.[37-39] Coherent oscillations in tubulin dimer subunits whose phases at discrete time intervals depend on lattice neighbor interactions may be considered as an automaton model of dynamic biological organization.

AUTOMATON MODEL OF MICROTUBULE INFORMATION PROCESSING

Dynamic patterns and information processing in automata depend on neighbor interaction rules among lattice subunits.[37-39]

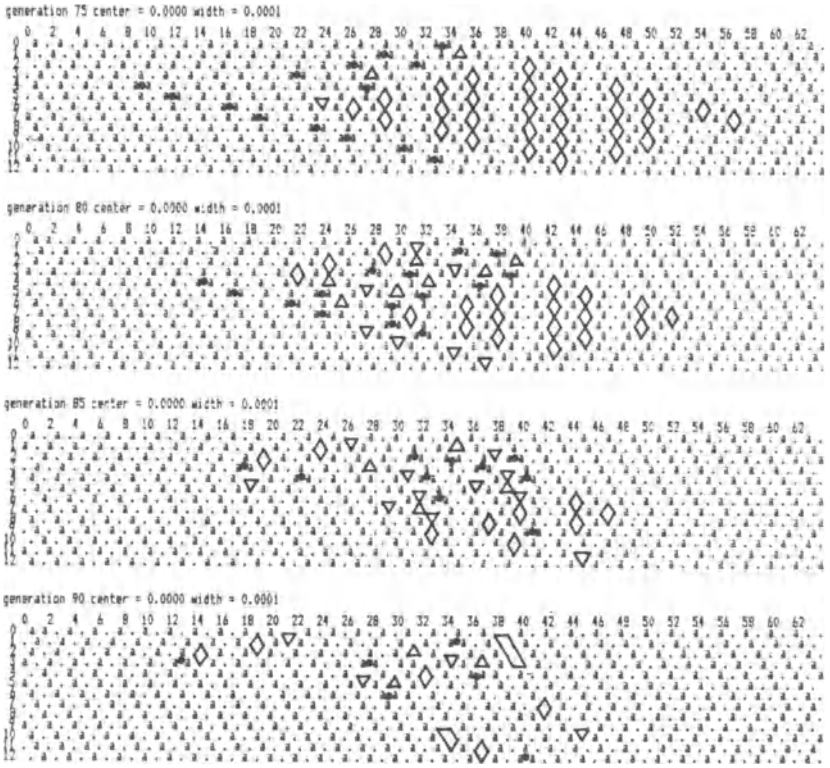

generation 75 center = 0.0000 width = 0.0001
generation 80 center = 0.0000 width = 0.0001
generation 85 center = 0.0000 width = 0.0001
generation 90 center = 0.0000 width = 0.0001

Fig. 8. Generations 75, 80, 85, and 90. Collision and interference of diamond patterns with triangle, parallelogram, and trapezoid structures.

Frohlich coherent excitations (10^{-10} sec) could effect discrete time intervals in MT automaton neighborhoods. Tubulin dimers as basic units would have a 7-member neighborhood on a hexagonal cylindrical lattice (Figure 4). The cylinder has a circumference of 13 units and the pitch pattern of the leftward helix would be one and one-half dimers (three tubulin monomers).

Each dimer may be viewed as a dipole with an electron shared by the two monomers. At each time-step (e.g., 10^{-10} sec), the electron's average position is in the center of one or the other monomer. The electrostatic force exerted on the electron by its nearest neighbors would depend on the neighbor states at the previous time interval, and may serve as "rules" for an automaton. The magnitude of the force exerted by any one neighbor is:

$$\frac{-k(e^2)}{r^2}$$

where k is $1/4\pi\varepsilon_0$, e is the charge on an electron and r is the distance between the electrons.

The location of the electron at any time can determine the conformational state of the dimer. We will assume each dimer is in either an alpha state or beta state at each 10^{-10} interval, depending on dipole orientation. The dimer will change state if the sum of the forces on its electron are sufficient to push it to the other side. The dimer neighbors are labeled with "north" (n) in the longitudinal direction ("y" axis, now horizontal) of the MT cylinder.

sw nw   -x ↑
s c n   y →
se ne

Only the relative magnitudes of the forces along the axis of the MT are important. Taking the MT axis to be the y direction, the force F(y) in the y direction on the electron in the center dimer is proportional to:

$$\Sigma \sin \theta / r^2$$

taken over all neighbors.

$$\sin \theta = y/r$$

therefore F(y) is proportional to y/r^3. Along the MT axis, tubulin subunits are 5nm apart, and around the circumference they are 4 nm apart. The helical twist yields a y offset of

$$4\text{nm}(3/13) = .9231\text{nm}$$

for nw and sw neighbors and -.9231nm for ne and se neighbors. MT automaton rules are shown in Table 1.

There are two parameters which can be varied in this model relating to the transition region where the force on the electron is near zero. The "center" provides a bias which prefers one state or the other due to, perhaps, protein structure, ion binding, or other factors. The "width" determines a neighborhood around the center where the force is not great enough to influence the next state. It is in some sense a measure of the "viscosity" of the system, the tendency for a state to remain stable (or oscillatory) until a certain threshold is reached. "Width" and "center" are both in units of y/r^3 or nm^{-2}. Forces less than the value of "center" minus "width" yield alpha states. The fate of forces within "width" of "center" is determined by the sign of "width." A negative width implies that a dimer will switch states regardless of the force on it (as long as the force is less than "width" from "center"). This implies that some other mechanism is causing

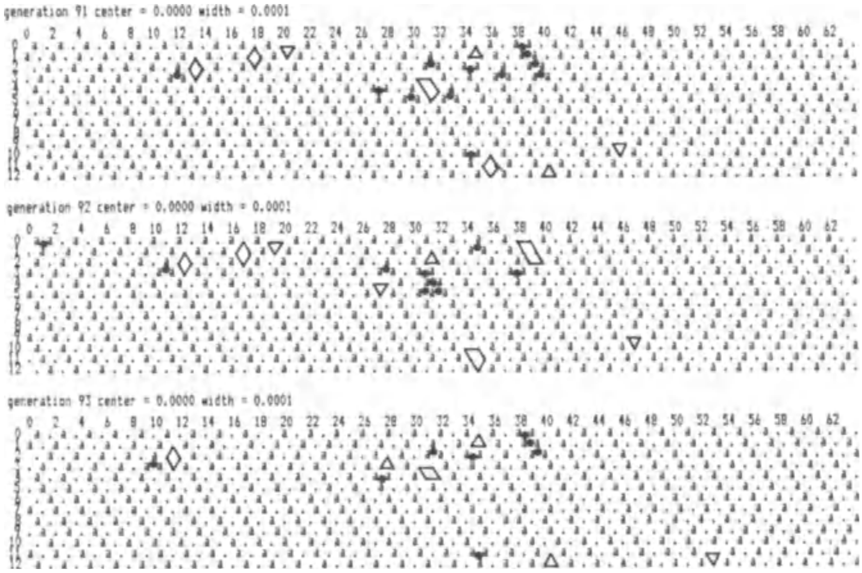


Fig. 9. Generations 91-93 with further interference and annihilation of structures and altered background patterns.

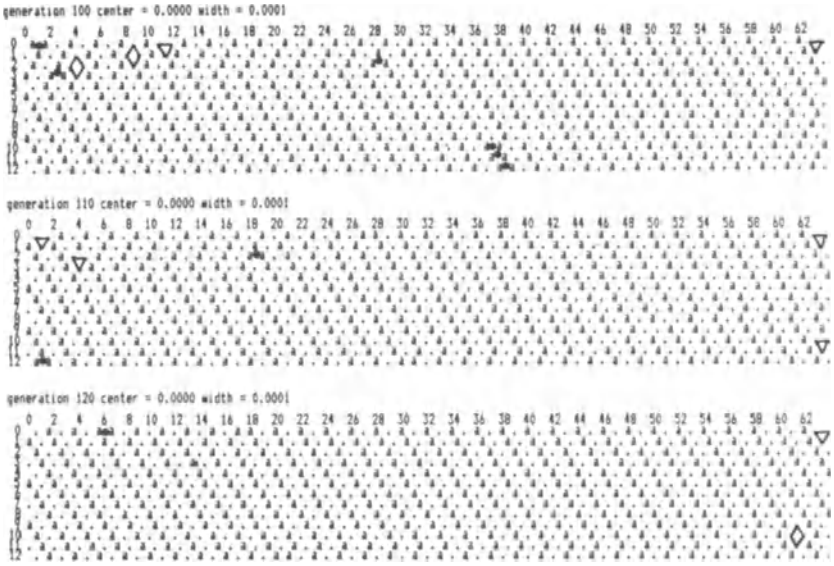


Fig. 10. Generations 100, 110, and 120. Minimal structure persists with altered background patterns as residue (? memory).

oscillation of the dipole. A positive width implies that for very small forces the state will remain the same. In the case of "center" at zero and a very small positive "width," several static, oscillatory and dynamic structures appear. All other variations of the rules will be treated as perturbations of the behavior of these rules.

In a field of dimers in the same state, occurrences of dimers in the opposite state "invade" the others until an approximately equal mixture exists. Figures 5-10 illustrate alpha states invading beta states. Except for the abnormal hexagonal geometry, the progress of the invasion is like the ripples from a stone dropped in a pond. The front wave advances steadily while standing waves oscillate behind. Since the rules are by nature quantized, any "width" (viscosity) less than the magnitude of the smallest possible unbalanced force will yield undamped oscillation. This means that wave fronts will travel forever and the standing waves left behind will continue to oscillate indefinitely. There appears to be a small set of structures which exist independent of initial conditions. Alpha and beta sites tend to alternate along the MT axis while they tend to align (a-a,b-b) along the (nw-se, sw-ne) axes. This yields rows of alternating bars which oscillate from alpha to beta and back. Disturbing this regularity can exist kinks and change of axis (nw to ne) which are also oscillatory with period two. The one and one-half pitch helix geometry prevents continuous bands longer than one revolution of the cylinder to occur in the sw-ne direction without a kink or direction reversal. In addition to kinks and change of direction, opposed pairs of triangles and other structures consisting of alpha and beta sites appear and oscillate also. These pairs also align themselves along the two off axes (ne-sw,se-nw).

The dynamic structures which exist in this environment appear to all be stable and/or interacting combinations of two basic structures. The simplest is just a pair of alpha or beta sites along the main axis. Alpha pairs travel south along the axis while beta pairs travel north. The other structure is a diamond arrangement, four alpha or beta sites surrounded by their opposite. These diamonds travel in opposite directions (alpha north, beta south). Both types of structures interact with their environment and can be destroyed by each other as well as absorbed by the stable triangle pairs. They also disturb the background pattern as they travel, reorganizing the structure as they go. These structures have also been observed to "bounce" off of each other or exchange phase. Figures 5-10 show boundary conditions causing reflection and initiation of patterns of diamonds and other structures which then interfere, resulting in altered patterns. These traveling patterns have velocities (8 nm/10^{-10} sec ≃ 80 m/sec) comparable to solitons and/or nerve conduction velocities, and may serve to transport binding sites or intracellular information.

CONCLUSION

Nonlinear electrodynamic phenomena including Davydov solitons and Frohlich coherent oscillations provide theoretical mechanisms for dynamic organization of biological activities. Cytoskeletal lattice proteins including microtubules are intimately involved in dynamic biological organization, and have structural geometries conducive to information processing by nonlinear electrodynamic phenomena. This paper has presented possible modes of microtubule information processing and dynamic organization. Frohlich coherent oscillations can result in automaton-like information patterns, propagating solitons, and regulation of a wide variety of dynamic cellular activities.

REFERENCES

1. A. F. Lawrence and W. R. Adey, Nonlinear wave mechanisms in interactions between excitable tissue and electromagnetic fields, Neurol Res 4:15 (1982).
2. A. S. Davydov, The migration of energy and electrons in biological systems, in: "Biology and quantum mechanics," A. V. Holden and B. Winlow, eds., Springer Verlag, Berlin (1982).
3. H. Frohlich, Coherence in biological systems, Collective Phenomena 3:139 (1981).
4. A. C. Scott, The laser-Raman spectrum of a Davydov soliton, Physics Letters 86A:60 (1981).
5. J. S. Clegg, Intracellular water, metabolism, and cellular architecture, Collective Phenomena 3:289 (1981)
6. S. R. Hameroff and R. C. Watt, Information processing in microtubules, J Theor Biol 98:549 (1982).
7. K. R. Porter and J. B. Tucker, The ground substance of the living cell, Sci Am 244:56 (1981).
8. P. Dustin, "Microtubules," Springer Verlag, Berlin (1978).
9. E. Lazarides, Intermediate filaments as mechanical integrators of cellular space, Nature 283:249 (1980).
10. J. Atema, Microtubule theory of sensory transduction, J Theor Biol 38:181 (1973).
11. K. Roberts and J.S. Hyams, "Microtubules," Academic Press, London (1979).
12. L. A. Amos and A. Klug, Arrangement of subunits in flagellar microtubules, J Cell Sci 14:523 (1974).
13. J. Bryan and L. Wilson, Are cytoplasmic microtubules heteropolymers? Proc Natl Acad Sci, 68:1762 (1971).
14. J. Bryan, Definition of three classes of binding sites in isolated microtubule crystals, Biochem 11:2611 (1972).
15. M. L. Shelanski, F. Gaskin, C. R. Cantor, Microtubule assembly in the absence of added nucleotides, Proc Natl Acad Sci 70:765 (1973).

16. R. F. Luduena and D. O. Woodward, Isolation and partial characterization of α and β tubulin from outer doublets of sea-urchin sperm and microtubules of chick embryo brain, Proc Natl Acad Sci 70:3594 (1973).
17. R. F. Luduena and D. O. Woodward, α and β tubulin: separation and partial sequence analysis, Ann NY Acad Sci 253:272 (1975).
18. G. G. Borisy and E. W. Taylor, The mechanism of action of colchicine: colchicine binding to sea urchin eggs and the mitotic apparatus, J Cell Biol 34:535 (1967).
19. H. Stebbings and C. Hunt, The nature of the clear zone around microtubules, Cell Tissue Res 227:609 (1982).
20. R. G. Burns, Spatial organization of the microtubule associated proteins of reassembled brain microtubules, J Ultrastruct Res 65:73 (1978).
21. J. C. Paulson and W. O McClure, Microtubules and axoplasmic transport, Brain Res 73:333 (1974).
22. S. Ochs and N. Ranish, Characteristics of the fast transport system in mammalian nerve fibers, Neurobiol 1:247 (1969).
23. M. Singer, Neurotrophic control of limb regeneration in newts, Ann NY Acad Sci 228:308 (1974).
24. M. Jacobs, Tubulin nucleotide reactions and their role in microtubule assembly and dissociation, Ann NY Acad Sci 253:562 (1975).
25. M. Jacobs and M. Caplow, Microtubular protein reaction with nucleotides, Biochem Biophys Res Commun 68:127 (1976).
26. S. P. Layne and A. C. Scott, An hypothesis of barbiturate action mediated via membrane and cytoskeletal proteins, in: "The Neurobiology of Pain," A. V. Holden and B. Winlow, eds., Springer Verlag, Berlin (in press).
27. J. M. Hyman, D. W. McLaughlin, and A. C. Scott, On Davydov's alpha-helix soliton, Physica 30:23 (1981).
28. A. S. Davydov, The theory of contractions of proteins under their excitation, J Theor Biol 38:559 (1973).
29. R. Weisenberg, The role of nucleotides in microtubule assembly, Ann NY Acad Sci 253:573 (1975).
30. R. C. Weisenberg, W. J. Deery, Role of nucleotide hydrolysis in microtubule assembly, Nature 263:792 (1976).
31. M. Ventilla, C. R. Cantor, and M. Shelanski, A circular dichroism study of microtubule protein, Biochemistry 11:1554 (1972).
32. H. Frohlich, Long range coherence and energy storage in biological systems, Int J Quant Chem II:641 (1968).
33. H. Frohlich, The extraordinary dielectric properties of biological materials and the action of enzymes, Proc Natl Acad Sci 72:4211 (1975).
34. H. Frolich, Long-range coherence and the action of enzymes, Nature 228:1093 (1970).

35. E. Del Giuduce, S. Doglia, and M. Milani, Self-focusing of Frohlich waves and cytoskeleton dynamics, Physics Letters 90A:104 (1982).
36. E. Del Giudice, S. Doglia, and M. Milani, Order and structures in living systems. This volume.
37. J. von Neumann, "Collected Works," A. H. Tareb, ed., 5:288 (1966).
38. M. Gardner, Mathematical Games, Sci Am 224:112 (1971).
39. S. Wolfram, Statistical mechanics in cellular automata, Caltech preprint CALT 68:915 (1982).

BRIDGING THE GAP

Samuel Koslov

The Johns Hopkins University
Applied Physics Laboratory
Laurel, Maryland

INTRODUCTION

We will review some of the historical background, biological and political, of the last two decades. The Pandora Project, the ERMAC Committee, the confusion engendered by different Federal Agency roles and admissions, issues involved in broadcasting regulation, all compounded by the public concern over the environment. Both the evolutionary development of man and our current knowledge of neurophysiology has long led to at least an intuitive belief that electromagnetic energy must influence body processes. However, until recent years there has been almost a total failure to confirm this other than in simple thermalization of the energy. The intrinsic complexity of biological systems has made it difficult to apply the rather simplistic theorems of the physical world. However, the increasing capability of the computer coupled with a mathematical structure of the behavior of physical systems with non-linear characteristics are beginning to lead to a breakthrough. An attempt will be made to summarize some of the key technical points that have been presented during the conference. Some recent results of a collaborative program at Johns Hopkins University permitting observation of a single cellular layer in vivo illustrating electromagnetic interaction phenomena will be summarized.

Integration consolidation and amplification in the governmental, fiscal, and technical domains are urgently and substantially required. A principal of "correspondence" parallel to that in quantum physics must be developed. The detailed processes we are beginning to uncover in the single cell must extrapolate to the megamouse experiment and eventually to the exposed human population. Only through such enhancement of our scientific knowledge base can credibility be achieved in the 3-way discussions between the researchers, regulators and the public. The role of the Federal Government to date in this critical field can be most kindly described as ranging from pathetic to abysmal. We are on the verge of a totally new area of understanding of biological processes, medical therapy and diagnosis, and rational control of the electromagnetic environment.

BACKGROUND

A concluding statement made in a different context might also serve as a summary to this past week.

"The biological action of electromagnetic fields most often amounts to some effect on the processes of control and intercommunication in the living organism: between systems, between cells, or between molecules. In other words, an electromagnetic field exerts an influence on the informational interactions in the organism, and apparently, the energy of the field serves only as a means of exerting this influence. In fact, it has been established experimentally that the nature of a given physiological reaction to an electromagnetic field remains practically invariant when the amount of energy in the acting field is varied over rather wide ranges. Furthermore, the effectiveness of the reaction can even increase with decreasing intensity of the treatment.

All this leads us to propose that the biological activity of electromagnetic fields is not due to energetic interaction, but to informational interaction with living organisms. That is, the major interaction here is not the transformation of electromagnetic energy into other forms, but the effect of electromagnetic fields on processes of transformation, transfer, coding, and storage of information in living systems. Further research will show how valid this conception is. (Ref. 1).

This statement was in the concluding section of an article published by A. S. Presman in "Fizika Nauk," June 1965. It indirectly had a stimulating affect on some programs we heard about today.

Research prior to 1965 (and to the present) could be categorized as either scientific or safety related. A noteworthy example of pre-1965 scientific work was that of Schwan and his co-workers on the electrical impedance of tissues. About 1967 concern over microwave ovens led to some enhancement of the safety program. Interest had been lagging for several years following the military safety programs engendered by the installation of the large surveillance radars of the late '50's and early '60's.

Two major scientific thrust were initiated by national security concerns. Pandora started in 1965, because of the unknown reasons for the beaming of microwaves at the Moscow Embassy, coupled with the somewhat obtuse (in terms of experimental data) statements by Presman and others in the Soviet literature with its heavy emphasis on low level research having no parallel in the West. There was a need to understand. Was there something we should know particularly about effects of EMR on the human central nervous system?

The program revolved around a principal experiment at the Walter Reed Forest Glen Annex. It was necessary to construct a totally new facility, since nothing suitable existed elsewhere. The radiation intensity used was an order of magnitude higher than the measured Moscow amplitudes. The equipment was capable of simulating the peculiar modulations actually observed.

The protocol looked for behavioral changes in operant conditioned chimpanzees. Many unwarranted artifactual conclusions were drawn during the course of the experiment. The net result was statistically inconclusive (Ref. 2). The intuitive tendency was to say nothing unusual had been observed.

The second area related to national security concerns started as a "safety" research program. The "Sanguine" (now Seafarer) system was intended to provide high reliability/low data rate communications with the strategic submarine forces using modulations on a carrier between 40 and 90Hz. An early biological study showed many drastic effects, which turned out rather eloquently to confirm Bob Frosch's aphorism "The best way to discover a new effect is to do a bad experiment." These early results and public opposition lead to a painstaking search for subtle affects at extra low frequencies (ELF).

Indirectly, another initially safety related activity contributed to the scientific development. A White House advisory committee, JTAC, prophetically suggested that safety problems both real and imaginary might inhibit civilian and military telecommunications in the future. This led to the formation of the Electromagnetic Radiation Management Advisory Council (ERMAC), advisory to Office of Telecommunications Management (OTM) - OTM became Office of Telecommunication Policy (OTP). Its director was also the Special Assistant to the President for Telecommunications. A principal OTP function was approval of frequency allocations for government users analogous to the FCC's role for commercial users. As Special Assistant he also had overall responsibility for national telecommunication policy. ERMAC assisted in the review of specific systems but soon realized that many unknowns, both scientific and pragmatic, existed. They began to plan a national program with a rate of growth built in (Ref. 3) and became a technical program coordinating group as well as a fiscal support point for 13 or 14 government agencies. Early on it was realized that a "sound scientific basis" was essential for standard setting. Without here detailing the philosophy of agency budget structures and actions, it was possible on White House stationery to either support requested funds or even force increases by means of OMB actions.

This all started grinding to a halt when the Carter administration transferred OTP responsibilities about 1978 to the newly formed National Telecommunications and Information Administration (NTIA) in the Department of Commerce. At present, the official NTIA position is that they have no responsibilities for this area. Even earlier, Department of Commerce letterheads didn't influence other agencies very much.

During a ten year period, the 1970's, the overwhelming percentage of funds came from military sources, greater than 60%, as much as 70 to 80% at times. This is no longer true. At no time have any substantial funds come from the pure science organizations such as NSF or NIH. NIH sporadically has inserted funds for the pragmatic study of thermal effects on cancer cells, but little for scientific investigation. Regulatory agencies, the Environmental Protection Agency (EPA), Bureau of Radiological Health (BRH), etc., have split their funds between the research and the regulatory aspects of their missions. EPA is concerned with exposure of the public, Occupational Safety and Health Administration (OSHA) with occupational exposure and NIOSH supports OSHA. BRH (now combined with medical devices) is concerned with control of emissions from electronic devices. The Federal Communication Commission (FCC) grants licenses for civilian spectrum use, while NTIA approves government usage. The Department of Defense is concerned with the protection of its personnel, including compliance with OSHA regulations (Ref. 4).

Two fundamental issues must be surfaced. First, there is no longer even an information or funding coordination point, as there was in the OTP-NTIA/ERMAC days; second, there is no agency with primary responsibility or even an interagency committee with defined responsibility for this area.

SOME HIGHLIGHTS OF THE MEETING

No attempt will be made to be either complete or comprehensive. However, a personal set of impressions of some highlights may be of value. During the last two decades three major directions have emerged in our understanding of living systems. The most visible is DNA structure and its manipulation, "recombinant DNA", bringing forth vast amounts of government and private efforts and funds with many applications. The second key thrust are the beginnings of a model for explaining the functional complexity of

*NIOSH - National Institute of Occupational Safety and Health the cell membrane with increasing applications, for example in immunology. Most recent has been the realization of the complexity of the structure of the extra-nuclear cell body and the beginnings of a model of the cytostructure. Taken together there has been an increasing intuitive feeling that indeed the "leg bone is connected to the thigh bone" and that the functions of each of these subsystems may be interdependent. There is no need here to belabor the problem of interdisciplinary collaboration between biological and physical scientists and the related difficulties of language. We can stipulate that these problems exist. What we have been talking about borrows from a number of fields of physical science in addition to the main thrusts of modern biology. Not surprising is that a number of useful concepts are derived from solid state physics, including refined classical and quantum models of dielectrics. Two major related concepts derive from somewhat more unexpected sources, from hydrodynamics and from mathematical theory. These are solitons and "chaos." Solitons were first described in observations of laminar flow in the mid-nineteenth century. In the mid-1950's Fermi, Pasta and Ulam studied the randomization of thermal energy in a computer model of a string of particles with non-linear coupling. Thermalization did not occur, rather a series of recurring, discrete, vibrational modes were observed. Similar behavior was found in the solution of the Kortweg-deVries equation, the "Granddaddy" equation for solitons (Ref. 5).

An exciting and very "hot" area of both mathematical and physical research today is generally subsumed under chaos theory. Some simple deterministic equations can exhibit surprising varieties of dynamic behavior, including stable points and stable cycles, in the presence of non-lineararities (Ref. 6,7). In effect resonance may occur for a given amplitude and frequency of a forcing function, disappearing when the forcing function amplitude either increases or decreases. These almost universal properties of many differential equations has surprisingly simplistic numerical relationships as given by the so called Feigenbaum numbers (Ref. 8). Up to now we have been somewhat trapped by the quantum theory. For example the feeling of the necessity for finding a quantum resonance. It now appears that there may be classical phenomena, universal at least within very broad classes, providing metastable states very sensitive to both the amplitude and frequency of the forcing functions.

Thus, two major somewhat related tools have appeared. The ability to predict oscillatory states by chaos theory and the means for propagation of a signal without or nearly without dissipation by a soliton. The possibility of soliton propagation modes, both through and along the cell membrane, and the existence of discrete "classical" energy states engendered by non-linear biological parameters may be the essential clues to the explanation of recently observed phenomena (Refs. 9,10,11). It's very early in this game. We are still responsible for meeting the requirements of scientific proof. Including both reproducibility and predictability. Reproducibility has been achieved in some cases, but quantitative predictability less so.

Much of the material presented at the conference falls into a few groups. A favorite subject seems to be the behavior of the neural membrane. Receptor activity including hormones, neuro transmitters, and anti-body action is another dominant theme. The mechanics of energy transfer, particularly solitons, as well as the excitation of, for want of a better term, resonance states, has been approached by several authors.

A number of different observational modalities are reported including Raman spectra, millimeter wave spectra and biological activity or biochemical assay. A number of different trends in theoretical development are apparent. These include solitons or other quasi particles, proton tunnelling, and cooperative (collective) models. Disciplinary domains drawn on include solid state physics, nonequilibium thermodynamics, nonlinear dynamics and the quantum mechanics of macromolecules in addition to those fields which constitute the forefront of modern biology. Overlapping all of these fields applied mathematics and computer "experiments" are increasingly significant.

Three characteristics of reported observations are recurrent and may be, together with theory, the key to understanding the mechanisms of electromagnetic interaction with biological systems. These are: frequency windows, amplitude windows and finite delay times between exposure and observable change.

RECENT EXPERIMENTS AT JOHNS HOPKINS

An on-going experiment at The Johns Hopkins University is collaborative between the Applied Physics Lab and the Medical Institutions. The principal investigator, Henry Kues, will present this in detail at the June 1983 Bioelectromagnetics Society meeting. It's significance to this conference is that it may represent a directly observable interaction between electromagnetic energy and a cell membrane. An observational technique not too well known in the research community (although becoming fairly standard in the clinical community) is used for direct in vivo observation of a single layer of cells. The program initially was started on internal R & D funds and is now supported by the Office of Naval Research. We sometimes call it "A funny thing happened on the way to the retina."

For many years Hopkins has had a team working on improved methods of retinal angiography. They thought it might be interesting to see whether the retina was affected by microwave radiation.

In order to assure that there were no optical artifacts in observing the retina, a recently available clinical instrument was brought into use to to see whether there were any corneal changes. (Prior studies, conducted under Army sponsorship, had shown gross corneal damage by infrared radiation (lasers) although at substantially higher levels.) The instrument used is known as a specular microscope. This instrument permits observation in vivo of the endothelial layer of the cornea. This is the rearmost single layer of cells which separate the cornea from the internal eye matter. One function of this cell layer is control of the hydration of the cornea. The specular microscope was first introduced into clinical practice in 1975 with a rather narrow field permitting the observation of some 50 to 150 cells at a time in the central four millimeters of the corneal endothelium. A year or two ago the wide-field specular microscope was introduced which can examine a square millimeter of endothelium, approximately 3,000 to 3,500 cells at one time and can scan the entire central 6 to 8 millimeters of the corneal endothelium (Fig. 1).

The initial experimental animal was the rabbit. There are some fundamental differences in the biology of the rabbit endothelium as opposed to the adult primate endothelium. In the rabbit if there is damage to the cornea, mitotic regeneration will occur. This mitotic regeneration has not been observed in the adult primate cornea, at least not in the central region. We might note that the clinical requirement of the development of the specular microscope was that damage to the endothelial layer, particularly following surgery, for example cataract surgery, is a major cause of blindness because of this failure to regenerate. If cells are damaged by surgical trauma for example, they will not regenerate. Megacells are produced to maintain continuity of the layer. If enough cells are lost, the transparency of the cornea is affected.

The focus here will be on the results of these experiments The details of protocol and apparatus will be published by the principal investigators in the near future (Ref. 12). Let me merely stipulate that the microwave exposure conditions have been very carefully and independently checked. A number of possible artifacts have been looked at in great detail, including, for example, insertion of a microthermistor directly into the cornea to measure temperature rise (always less than a degree), and test exposures have been carried out with eye lids stitched to minimize any question of dehydration of the corneal surface. A standard mild anesthetic regimen used in many other experiments at the Hopkins School of Medicine was employed. A full set of controls, including shams were intrinsic to the experiment.

The results of the initial experiment on rabbits were that CW radiation at 2450 Mhz produced corneal changes with distinct

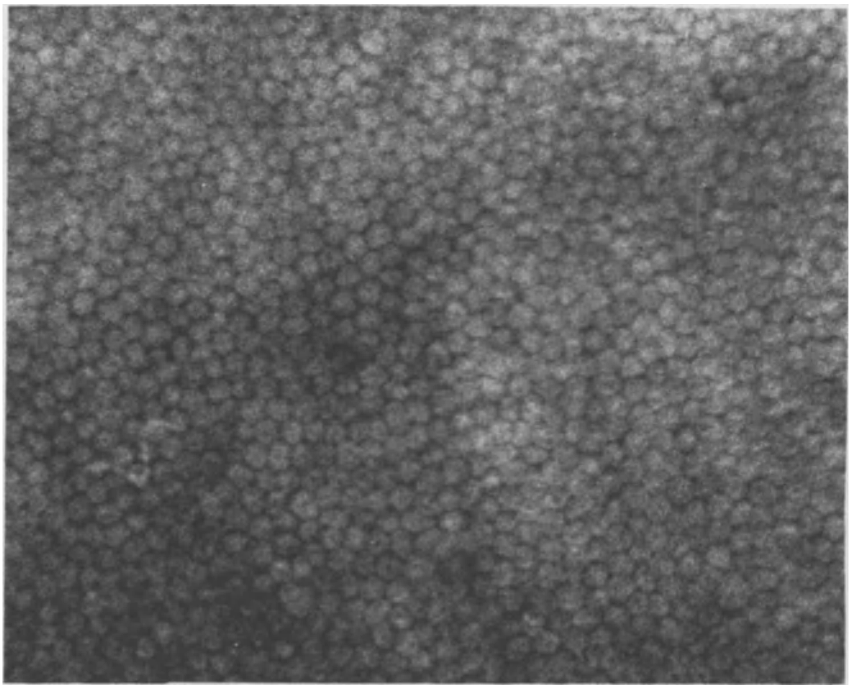


Fig. 1. Specular photomicrograph of the normal corneal endothelium of cynomolgus monkey.

cell pathology after a single four hour exposure at 10 or 20 mW/cm^2. Obviously, well below the accepted level for cataractogensis. An intriguing feature of these observations was that the cell damage was not observed immediately after exposure. Rather, it took some 24 to 48 hours for the cellular pertubations to be seen. The focus of the experiment was then transferred to primates. Cynomolgus monkeys, comparable in size to Rhesus, were used. No abnormalities were seen in sham exposures. At 30 mW/cm^2, abnormalities were seen in all seven primates examined. One showed them at 10 mW/cm^2. The same time delays in the appearance of the abnormalties were observed.

The waveform was changed to a pulse mode maintaining an average power of 10 mW/cm^2, the lower value at which changes had been observed. The standard regime was a 10 microsecond pulse, 100 pulses per second. The peak power thus was 1,000 times greater than the average of 10 mW/cm^2. In 11 experiments, 8 showed abnormalities (Fig. 2). More recent experiments have shown that pilocarpine, a topical drug commonly used in the treatment of glaucoma, tends

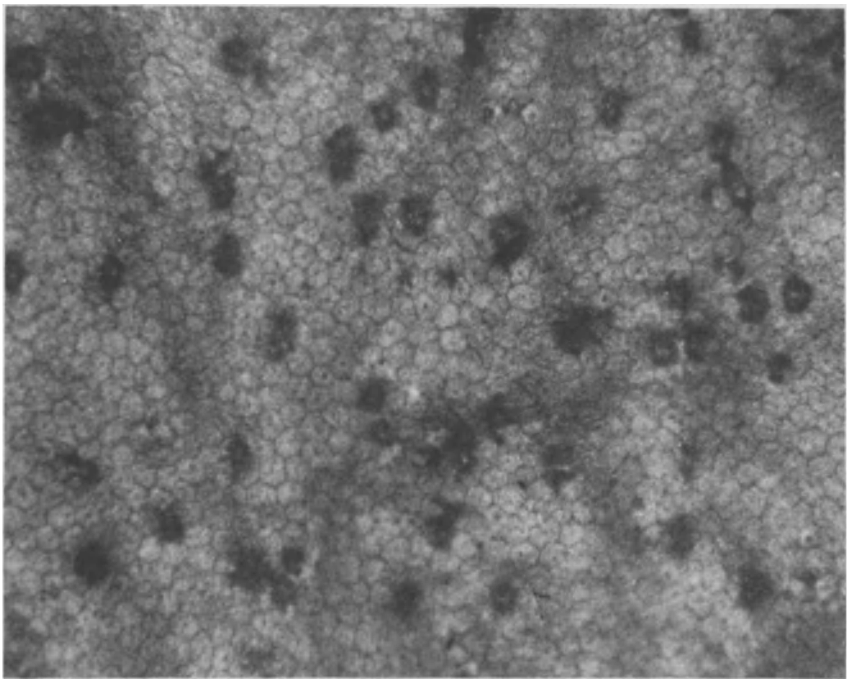


Fig. 2. Typical appearance of the corneal endothelium 48 hours after a single 4 hr. exposure to 10 mW/cm^2 (avg.) (2,450 MHz) pulsed microwaves.

to reduce the threshold for the effect. Initial experiments seem to show that the effect will occur at an average power of 5 mW/cm^2 (pulsed) when this drug is used.

The observational techniques extend beyond the specular microscope. Both detailed optical (Fig. 3) and electron microscope (EM) histology (Fig. 4) have been used. On-going experiments are using radioactively labeled glutathione and thymidine in order to elucidate mechanisms. The synergistic action of the pilocarpine is rather intriguing, since its pharmacological properties are fairly well established. In vivo investigations ongoing relate primarily to the endothelial "pump" function, which is affected by the pilocarpine.

In summary, some points of interest here are: First, the observation of cellular change at levels well below the so called "thermal" level (10 mW/cm^2 corresponds to a SAR of about 2.6 W/kg). Secondly, the effect does not occur immediately after radiation

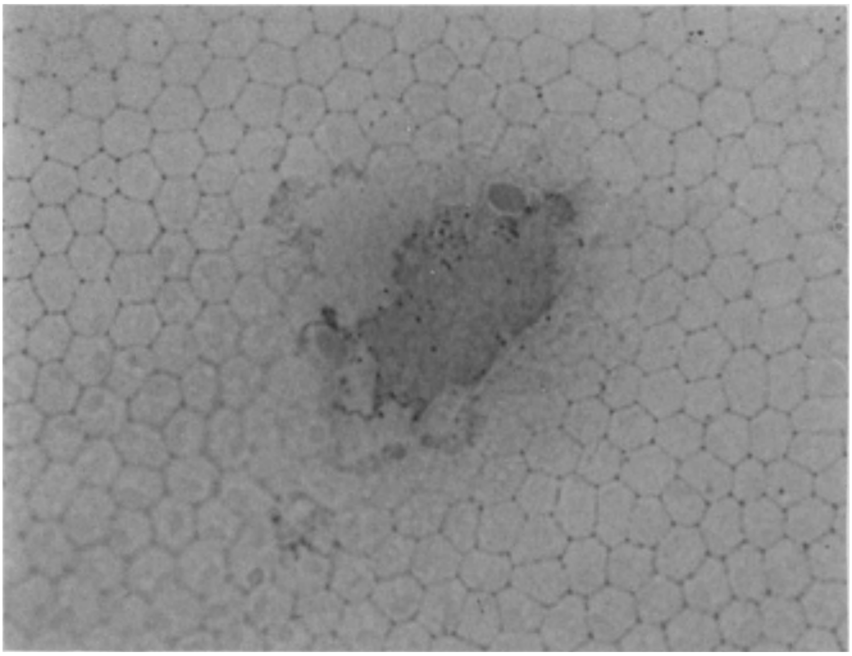


Fig. 3. Photomicrograph of excised corneal endothelium (Alizarin red/Trypan Blue stain) showing cell trauma.

but rather has a latency period that appears to extend somewhere between 16 and 48 hours. Third, and perhaps most significant, the pulsed radiation is significantly more effective than cw radiation for the same average power.

From a safety viewpoint, these results are very interesting. Particularly in view of the rather generous allowance for exposure of the new ANSI standards at levels greater than 300 MHz and additionally in view of the rather common use of drugs like pilocarpine in the treatment of glaucoma among the public. From the scientific viewpoint, we believe we have a technique of looking at single cells in vivo under exposure with effects occurring that are related to the wave form. Investigations are continuing with plans to look at different frequencies and different modulations as well as to extend the number of observational modalities.

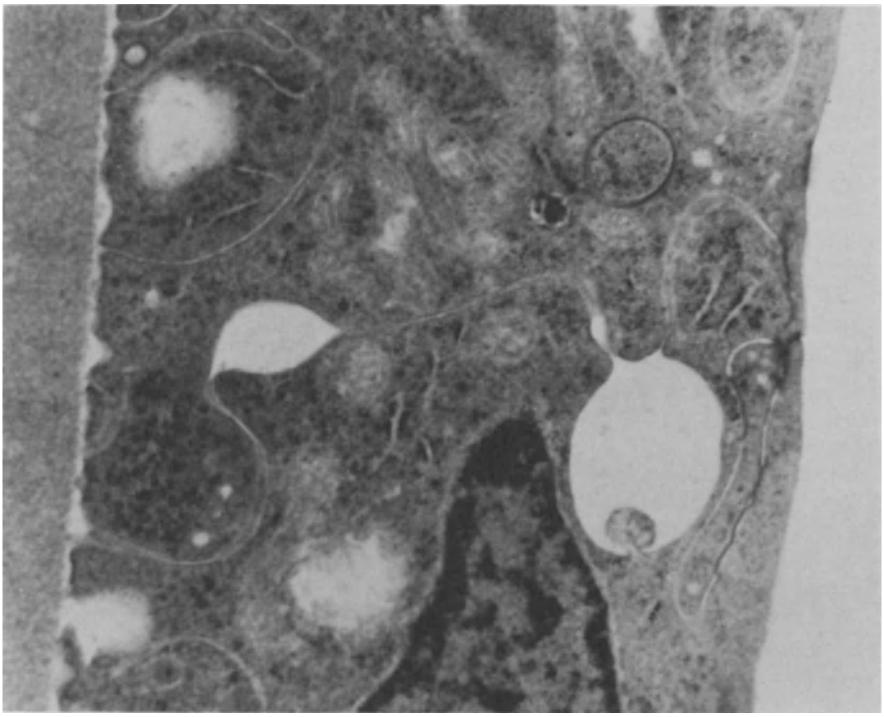


Fig. 4. Electron photomicrograph of a cross section of the endothelial layer following exposure. Note large vacuoles at cell interface.

RECOMMENDATIONS

The observations and hypotheses that have been presented during the conference are converging to a level of understanding of cellular behavior that implies the external electric fields can, in fact, become a key to the cellular control console. The implications, social, economic, and even military are enormous. However, the barriers to major acceleration of these efforts are also enormous. What we have heard requires unconventional thinking. It requires the academician to leave the warm, comfortable environment of his own discipline. It requires many of those who control the flow of funds both as peer advisors and as managers to admit that they didn't foresee a really significant direction of scientific understanding. In the present environment of budgetary restrictions, with increasingly larger areas of protective turf being defined thereby, simply going through the usual channels is likely to have little impact and will be fortunate if it can even maintain level of funding now extant for the areas we have discussed.

Before any Herculean effort can be made to substantially increase these activities, we must be sure that we can show proof of scientific reality, reproducibility and predictability. Predictability will, of course, will subsume many potential applications. If much of what we have heard is indeed correct, it may be no less significant to the nation than the prospects that faced the physics community in 1939 when the long-time predicted fissionability of the nucleus was actually demonstrated. You may recall the famous letter of Albert Einstein to President Roosevelt. When we're in a position to do so in terms of our proofs, I would propose that an analogous letter is required.

To start with, a small group is needed to reduce to an easily comprehensible size both the proof and its significance. Additionally, this small group must massage and refine a proposed plan for research and related requirements for facilities. I think a self organized committee must be formed to produce this "white paper." Such a committee, however, should not be drawn exclusively from those here at this conference. Rather it should consist of approximately half of its members drawn from the community here represented and half drawn from a more general scientific community, physical and biological and perhaps engineering, who are essentially virginal in so far as these research areas are concerned. This would serve two functions. First, to insure that indeed a statement can be made which is comprehensible to those who have not been working in the field. Second, to give an unbiased view as to whether, indeed, the proof has been established and prediction is possible.

An individual has to be identified who could be, but not necessarily has to be, a member of this group. The credentials and stature of this individual, be he a biological or physical scientist, must be such as to give him access to the highest levels of government, both executive and congressional. He will have to be both the signer and bearer of this letter, or its equivalent. The same individual may be appropriate, or a second has to be identified whose credentials are equally valid and significant to the industrial and the university communities. A similar approach must be then be taken with both industry and university.

Three things are necessary if there is to be a significant government sponsored program. Unless all three are extant and parallel, survival and maturing of the program is unlikely to occur. The three items are:

- An identified central point in the government; an agency or special committee that is the focal point for funding decision making.

- One or preferably more centers of excellence identified for development of this field, probably, but not necessarily, university connected. They could be associated with a government organization, such as DoD, VA or NIH.

- Third, and most important, a pool of funding available which is not committed to these national centers of excellence. Rather it must permit flexible decisions in funding a new approach, an outstanding, previously unidentified individual or a new organization. Failure to do this and preserve this ability for independent decision outside of the national centers will inevitably result, as it has in many other cases, in a large but stagnant pool of talent working forever toward goals they had thought of decades ago.

In parallel, or perhaps sequential, to the development of a government program, a collaborative industry/university program should be established. For the case of recombinant DNA many examples now exist. Examples abound in many other technology areas of potential profitability. These collaborations can take at least two forms or some combination thereof of these two forms. One type it is a knowledge sharing association whereby the results of research are shared with participants in the joint organization and where the industrial participants can use the facilities of the university including temporary interchange of their own people and the university people. A less palatable but nevertheless perhaps more effective scheme is where the industrial financial supporter receives priority in anything which may have patent related profitability. The possible commercial implication of the areas discussed here are enormous, ranging from pharmacology to biomolecular electronics.

These proposals may appear radical. However, simply encouraging everyone to sit down and write a proposal to their potential sponsors in the routine manner will only absorb energy that could be well used in scientific investigation with minimal end results. Some gain yes, to be sure. I believe we are on the verge of a total new understanding of the living system and I believe that we are obligated to do something about accelerating its applications.

REFERENCES

1. A. S. Presman, The Action of Microwaves on Living Organisms and Biological Structures, Soviet Physics Uspekhi, Vol. 8, No. 3, pp. 463-488 (Nov-Dec 1965).
2. "Unclassified Documents Pertaining to Project Pandora," Hearings before the Committee on Commerce, Science, and Transportation, United States Senate, 95th Congress, First Session on Oversight of Radiation Health and Safety, June 16, 17, 27, 28 and 29, 1977, Serial No. 95-49, U. S. Government Printing Office, Washington, DC (1977).
3. Report on "Program for Control of Electromagnetic Pollution of the Environment: The Assessment of Biological Hazards of Nonionizing Electromagnetic Radiation," issued by the Office of Telecommunications Policy (March 1973).
4. "Fifth Report on Program for Control of Electromagnetic Polution of the Environment - The Assessment of Biological Hazards of Nonionizing Electromagnetic Radiation." NTIA Report 79-19, U.S. Department of Commerce, National Telecommunications and Information Administration, Washington, DC (March 1979).
5. A. C. Scott, F. Y. F. Chu and D. W. McLaughlin, The Soliton: A New Concept in Applied Science, Proc. of the IEEE, Vol. 61, No. 10, pp. 1443-1483 (1973).
6. R. M. May, Simple mathematical models with very complicated dynamics, Nature, 261:459-467 (1976).
7. B. V. Chirkikov, A Universal Instability of Many-Dimensional Oscillator Systems, Physics Reports, 52, No. 5 (1979) 263-379, North-Holland Publishing Co.
8. M. J. Feigenbaum, Universal Behavior in Nonlinear Systems, Los Alamos Science, 4-27 (1980).
9. A. S. Davydov, Solitons in Molecular Systems, Physica Scripta, Vol. 20, 387-394 (1979).
10. Leonard S. Taylor, The Mechanisms of Athermal Microwave Biological Effects, Electromagnetics 2:259-267 (1981).
11. A. F. Lawrence and W. Ross Adey, Nonlinear Wave Mechanisms in Interactions between Excitable Tissue and Electromagnetic Fields, Neurological Research, Vol. 4, No. 1/2 (1982).
12. H. A. Kues, L. W. Hirst, S. A. D'Anna and G. Dunkelberger, Effects of 2.45 GHz Microwaves on the Primate Corneal Endothelium, to be published in Bioelectromagnetics.

ACKNOWLEDGMENT

This work was supported, in part, by the Department of the Navy, Naval Sea Systems Command, under Contract N00024-83-C-5301.

INDEX

Acetanilide, 149
  simulation, 151
Acoustic phonons, 211
Amide-I bonds, 144
  energy propagation, 136, 144, 147
  interactions, 146, 148, 152
  vibrations, 135
Amide spines, 531
  in anesthesia, 534
  in Davydov solitons, 532
  in helical proteins, 532
  soliton models of anesthesia, 540, 544
ATP hydrolysis, 134
Attractor, 403
Automata
  microtubules, 574
  microtubule lattice interactions, 576
  nearest neighbor states, 579
  pattern generation, 578

Bacteriorhodopsin, 107-120 (see also chromophore)
  dark reaction, 117-118
  proton pumping cycle, 108
  proton translocation, 107
  trans-cis photoisomerization, 114
  two channel pumping model, 113
Barrier, 227-228
  energy characteristics, 236-237
  Schottky, 236
Bifurcation, 465
Biological energy harvesting, 245
Bloch-Maxwell equations, 434
Bone cells, 23-33
  adenylate cyclase activation, 10
  cholera toxin action, 29-30
  cyclic-AMP production, 24-26
  magnetic field effects, 10
  parathyroid hormone response, 23-24
  regulatory N protein, 32
Boson condensation, 469-475

Calcium efflux, 6
Cell membrane, 3
  glycoproteins, 5
  nonlinear electrodynamics, 15
  solitons, 17
  transductive coupling, 16-17
Cell oscillations, 87
  Frohlich model, 88-89
  surface charge density, 90
Cell spin resonance, 87, 95
Chaos, 403, 408, 421, 424, 433, 437, 451
  biological relevance, 410
  deterministic, 441
  field-induced, 413
Chemical reactions, 277
  modulation effects, 281
  RF and product formation rate, 280

RF radiation, 277
Chromophore, 107
  counterion, 111, 118
  isomerization, 112, 118
  molecular model, 112
  resonance structure, 118
  Schiff base, 111
Coherent excitations, 433
  491-496
  limit cycles, 491, 495
  metastable polar states, 491, 493
  single mode, 491-492
  stabilized, 492
Collective behavior, 469-470
  dissipativity, 470
  electrets, 472-474
  entropy of living systems, 470
  Lagrangian function, 472-473
  quantum field theory, 470-471
  quasiparticles, 470
Computer technology, 207-218, 243-274
  acoustic phonons, 211
  device size, 209
  Josephson effect, 213
  molecular gates, 210
  organic polymers, 209
  submicron structures, 207-209
  thin films, 210
  tunneling, 213
Conducting polymers, 246
Content-addressable memory, 267
Control groups in molecules, 249
Corneal endothelium, 593
  endothelial damage, 593
  microwave effects, 594
  pilocarpine synergy, 594
Cytoskeleton, 567-583
  information processing, 577
  dipole oscillations, 574
  microtubules, 568
  nonlinear electrodynamics, 567
  protein lattices, 567
  solitons, 573
  tubulin dimers, 580
  tubulin subunits, 573

Dendritic potentials, 309
Detectors
  circuit, 232
  logic, 233
Dielectric properties, biological, 491
Dielectrophoresis, 92
Dipole-dipole interactions, 145, 147
Dipole radiation, 342

Electrets, 471-473
Electromagnetic field
  homogeneous wave, 340
  inhomogeneous wave, 340
  Norton ground wave, 341
  penetration, 339
  Zennick wave, 340-342
Electron-phonon interactions, 210
Electron tunneling, 227, 231, 247
Extracellular space, 307-310
  boundary conditions, 323
  Debye layers, 310
  dendritic potentials, 309
  field potentials, cerebral, 307
  information transfer, 313
  ionic oscillations, 321
  ionic resonances, 316
  neuron-glial relations, 312-313, 320
  neuronal populations, 307-308
  potassium ion movement, 312

Fermi-Pasta-Ulam lattice, 451
Field penetration, 340-342
Fixed points, stable, 436
Franck-Condon principle, 135

Hamiltonian
  annihilation operator, 135
  classical, 147
  creation operator, 135
  operators, 145
  perturbed two-particle, 415
  tight-binding, 201
Hydrogen bond, 144
Hydrogen-bonded chain, 112
Infinite dimensional map, 435
Insulin secretion in fields, 15
Integral proteins, 229
Interdisciplinary research
  U.S. Federal funding, 588
  U.S. Federal oversight, 587
Ionic dynamics
  in extracellular space, 312
  potassium resonance, 313

Josephson junction, 7, 8, 212-213
  critical current, 223
  switching energy, 224
  switching time, 224
  terahertz oscillators, 222

Langmuir-Blodgett films, 219-226, 230, 265
  band-gap, 222
  barriers, 220
  insulators, 220
  lithography, 221
  preparation, 220
  process compatibility, 224
  resists, 221
  thickness, 222-223

Limit cycle, 394
  biological relevance, 409
  collapse, 400
  entrainment, 398
  Lotka-Volterra model, 494
  microwave biological effects, 494
  onset, 395
  perturbation, 397
Lithography, 221, 260
  epitaxial deposition, 262
  molecular level, 260
Logic devices, 229
  biological, 229, 231-232
  Josephson junction, 229
  semiconductor, 229
Longitudinal sound waves, 135
Lossy medium, 339
Lymphocytes
  allogeneic cytotoxicity, 10
  protein kinase activity, 11
Lysozyme, 147
  simulation, 151, 153

Magnetic ring current, 342
Membrane model, 519
  diffusion in, 524
  transductive coupling, 16-17
Microscope
  specular, 591
Microwave fields, 6, 59, 475
  cerebral calcium efflux, 7
  lymphocytes in, 11
  modulation frequency, 11
  S-band, 339
Millimeter microwaves, 178-186, 549-566
  cavity, high-Q, 178
  coherent sources, 178
  desorption isotherms, 551
  hemoglobin spectrum, 179
  lysozyme spectrum, 177
  nonlinear phenomena, 549
  nonthermal responses, 552
  oncology, 549
  phonon processes, 184
  Raman spectroscopy, 553-554
  relaxation processes, 183-184
  spectroscopy, 178
Molecular electronic devices, 243
  chemical subunits, 244, 264-270

  hybrids, 244, 272-273
  Merrifield synthesis, 244-245
  molecular lithography, 244, 262
  prosthetics, 244, 273

Nerve bundles
  lobster, 520, 523
  optical scattering, 520, 523
  Raman spectroscopy, 521, 525
Nervous tissue, 65-77
  astrocytes, 69
  Bergmann glia, 69
  brain proteins
    in electric fields, 69
  endoplasmic reticulum, 68
  Purkinje cells, 68
Neurotransmitter
  calcium-dependent release, 84
  noradrenaline release in H-field, 81-83
  phaeochromocytoma cells, 81
  release in pulsed magnetic fields, 79-86
  release of GABA in pulsed EM field, 9
Nonlinear electrodynamics, 347
Nonlinearities
  models of
    amplifiers, 356
    electromagnetic, 347
    optical bistability, 371
    quantum models, 376
    radiation damping, 356
    selfinduced transparency, 352
    solitary waves, 355
    tetracritical behavior, 361
Nonlinear interactions, 143, 492
  longrange, 493, 495

Optical ring cavity, 433
Optical scattering, 522
Oscillations
  coherent macromolecular, 564
Oscillators
  coherent, 501
  van der Pohl, 501

Period doubling, 403-404
Period doubling sequence, 436
Permittivity
  dielectric, 340
Phase space
  flow, 418
  plots, 416
  trajectory, 416
Phonons, 140
Photocycle
  dark reaction, 116
  efficiency, 112
  energetics, 108
  intermediates, 109
  primary current, 116, 114
  time constants, 113
Phthallocyanine, 248-249
Polyacetylene, 201
Protein alpha helix, 133
Protein kinase activity, 11
Proteins
  picosecond relaxations, 177
Proton
  channels, 112
  deprotonation, 110
  electrochemical gradient, 108
  motion, 147
  protonation, 110
  pumping, 107-120
Proton transfer, 254
Proton tunneling, 259
Pulse propagation, 401
  onset, 401
  transition, 402
Purple membrane, 108

Quartic potential, 420
Quasiparticles, 470
Quasiperiodicity, 394

Raman spectroscopy, 554
  in cell division, 557
  harmonic generation, 562
  solitons in DNA, 562
  solitons in protein, 562
  soliton propagation, 557
Raman spectra, 526
Resonance
  nonlinear, 522
  zones, 523
Resonance structure
  bacteriorhodopsin, 117
Resonant transfer, 134
Recurrence, 451
RNA, 526

Schrodinger equation, nonlinear, 434, 437 (see also soliton envelope)
Self-synthesis, 244
Self-organization, 244, 264-270
Soliton, 17, 394
  activation, 255
  charged, 202
  Davydov, 133, 144
  envelope, 449, 455
  free energy, 138, 141, 144
  generation, 255, 260, 263
  kink, 201, 202
  lifetime, 141
  photoactivated, 254
  propagation, 255
  proton, 133, 135
  spin, 202
  switching, 252-253
  threshold, 138
  Toda lattice, 414
  valves, 258
  wave function, 136
Spectroscopy
  autocorrelation, 519, 522
  infrared, 519
  Raman, 519, 525, 554-562
Switching, molecular
  bias, 579
  element, 227

Time scales, 148
Tissue growth, 287
  cancer growth, 288
  dissipative structures, 287
  environmental effects, 289
  immune attack on tumors, 292
  reaction-diffusion, 287
  steady-state solitons, 293
  tissue-phase effects, 304
  traveling wave solutions, 294
Toda lattice, 414
  two particle, 415
Tunnel barriers, 219, 223 (see also electron tunneling, proton tunneling)
  characteristics, 223
  height, 224
  inelastic electron tunneling, 211
  Josephson junction, 219, 223
  MISFET, 222

Water waves, 443
  governing equations, 444
  linearized equations, 445
  long-time evolution, 451
  modulational instability, 448, 455
Wave trains, 445
  recurrence, 457
  stable mode, 451
  two space dimensions, 452
Windows
  frequency, 6, 10, 11, 398
  intensity, 6, 398

Yeast growth, 59, 61
  resonant effects, 64
  nonthermal millimeter wave effect, 60

Zakharov's integral equation, 457